



**UNIVERSIDADE FEDERAL DO CEARÁ**  
**CAMPUS DE RUSSAS**  
**CURSO DE GRADUAÇÃO EM ENGENHARIA DE SOFTWARE**

**RENÊ MICHEL SANTANA PADILHA**

**AUTOMAÇÃO DA IDENTIFICAÇÃO DE  $C_3S$  NA MICROSCOPIA DO CLÍNQUER**

**RUSSAS**

**2023**

RENÊ MICHEL SANTANA PADILHA

AUTOMAÇÃO DA IDENTIFICAÇÃO DE  $C_3S$  NA MICROSCOPIA DO CLÍNQUER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Software do Campus de Russas da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Software.

Orientadora: Prof. Dra. Rosineide Fernando da Paz

RUSSAS

2023

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas  
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

P134a Padilha, Renê Michel Santana.  
Automação da Identificação de C3S na Microscopia do Clínquer / Renê Michel Santana Padilha. – 2023.  
65 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Campus de Russas,  
Curso de Engenharia de Software, Russas, 2023.  
Orientação: Prof. Dr. Rosineide Fernando da Paz.

1. Clínquer. 2. Aprendizado de Máquina. 3. Automação. 4. Visão computacional. 5. Redes Neurais. I.  
Título.

CDD 005.1

---

RENÊ MICHEL SANTANA PADILHA

AUTOMAÇÃO DA IDENTIFICAÇÃO DE  $C_3S$  NA MICROSCOPIA DO CLÍNQUER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Software do Campus de Russas da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Engenharia de Software.

Aprovada em: 12 de Dezembro de 2022

BANCA EXAMINADORA

---

Prof. Dra. Rosineide Fernando da Paz (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Alexandre Matos Arruda  
Universidade Federal Do Ceará (UFC)

---

Prof. Dr. Yuri Lenon Barbosa Nogueira  
Universidade Federal Do Ceará (UFC)

## RESUMO

O cimento portland é o nome formal para o material comumente usado na construção civil. O clínquer é o cimento em sua penúltima etapa da produção, nessa fase da produção são comumente utilizadas técnicas para aferir sua qualidade. A análise do clínquer gera informações da qualidade do produto final, assim como do processo de produção. Esta análise pode ser feita por difração de raios X e/ou microscopia, sendo a última a mais utilizada. Mesmo com sua grande importância na produção de cimento, a microscopia é realizada de forma manual, o que pode levar a demora na identificação de problemas no processo produtivo e grande variabilidade nos resultados das análises. Tendo em vista o grande avanço na área de aprendizado de máquinas, é relevante considerar técnicas para automação dessa tarefa. O aprendizado de máquina é um método de análise de dados que automatiza a construção de modelos analíticos capazes de identificar padrões e, com isso, ter condições de otimizar os acertos em tomadas de decisões. Um tipo de algoritmo dentre os mais avançados na área de aprendizado de máquina são as redes neurais artificiais. Estes são modelos computacionais que funcionam de modo similar ao cérebro biológico. Assim, este trabalho tem por objetivo utilizar o algoritmo Mask R-CNN (do inglês *Mask Region Convolutional Neural Network*) (HE *et al.*, 2017) para a segmentação e classificação de cristais de  $C_3S$  em imagens microscópicas do clínquer. Uma vez que a forma e a distribuição desses cristais nas imagens fornecem evidências sobre a qualidade do clínquer produzido, a segmentação e classificação dos mesmos irá possibilitar que as análises que dependem de sua forma, tamanho de distribuição sejam automatizadas. Com a Mask R-CNN foi possível atingir resultados satisfatórios, porém, cristais de  $C_3S$  muito unidos entre si (com uma divisão muito discreta entre os cristais próximos uns dos outros) não foram separados (contabilizados como um só ou não foi identificado pela rede neural). Isso se deve ao fato de grande parte das imagens na base de dados são de qualidade inferior, somado a isso, o especialista em análise de clínquer não esteve presente durante o presente estudo afim de informar o que estava errado ou correto nos resultados de cada treinamento da Mask R-CNN.

**Palavras-chave:** microscopia do clínquer; automação na indústria cimenteira; segmentação de imagens microscópicas

## ABSTRACT

Portland cement is the formal name for the material commonly used in construction. Clinker is cement in its penultimate stage of production, at this stage of production techniques are commonly used to assess its quality. Clinker analysis generates information on the quality of the final product, as well as on the production process. This analysis can be done by X-ray diffraction and/or microscopy, the latter being the most used. Even with its great importance in cement production, microscopy is performed manually, which can lead to a delay in identifying problems in the production process and great variability in the results of the analysis. In view of the great advance in the area of machine learning, it is relevant to consider techniques for automating this task. Machine learning is a data analysis method that automates the construction of analytical models capable of identifying patterns and, therefore, being able to optimize the successes in decision making. One of the most advanced algorithms in the area of machine learning is artificial neural networks. These are computer models have inspiration of the operation of biologic neurons. Thus, this work aims to use the Mask Region Convolutional Neural Network (Mask R-CNN) algorithm (HE *et al.*, 2017) for the segmentation and classification of  $C_3S$  crystals in clinker microscopic images. Since the shape and distribution of these crystals in the images provide evidence about the quality of the clinker produced, the segmentation and classification of these crystals will enable analyzes that depend on their shape, size and distribution to be automated. With the Mask R-CNN, it was possible to achieve satisfactory results, however, crystals of  $C_3S$  very close together (with a very discreet division between the crystals close to each other) were not separated (counted as one or not identified by the neural network). This is due to the fact that most of the images in the database are of inferior quality, in addition to this, the clinker analysis specialist was not present during the present study in order to inform what was wrong or correct in the results of each training by Mask R-CNN.

**Keywords:** clinker microscopy; automation in the cement industry; microscopic image segmentation

## LISTA DE FIGURAS

Figura 1 – Exemplo de uma imagem de clínquer . . . . .	13
Figura 2 – Exemplo de um cristal de um poro e um cristal de alita . . . . .	14
Figura 3 – Exemplo de um perceptron de uma camada. . . . .	24
Figura 4 – Principais funções de ativação e seus respectivos comportamentos em gráfico. . . . .	25
Figura 5 – Exemplo de pontos de máximo e mínimo que podem ser encontrados na descida do gradiente. . . . .	27
Figura 6 – Exemplo de uma rede neural multicamadas com três camadas ocultas. . . . .	28
Figura 7 – Exemplo do fluxo de informações da retropropagação no interior da rede. Começando dos neurônios de saída até os de entrada. . . . .	28
Figura 8 – Ilustração da aplicação de um filtro convolucional. . . . .	30
Figura 9 – Ilustração do molde exibindo duas lamínulas construídas para obtenção de imagens para a realização da microscopia. . . . .	33
Figura 10 – Ilustração da imagem microscópica após ataque com amônia por 35 segundos. . . . .	34
Figura 11 – Diagrama ilustrativo para as etapas de processamento realizados nas imagens de clínquer até a fase de classificação dos componentes mineralógicos. . . . .	36
Figura 12 – Imagem de clínquer rico em $C_3S$ (cristais azulados). . . . .	37
Figura 13 – Comparação da aplicação do K-Means com a imagem normalizada e não normalizada. A mais à direita é a original sem processamento. . . . .	38
Figura 14 – Gráfico de distorção mostrando a performance do K-Means de acordo com seu parâmetro de quantidade de centróides. . . . .	39
Figura 15 – Segmentação executada pelo KNN. À direita são os poros presentes na imagem. . . . .	39
Figura 16 – Segmentações realizadas pelo K-Means e KNN. . . . .	40
Figura 17 – Exemplo do aumento de uma imagem de Magnésio, no centro(destacado em vermelho) o seu tamanho original 10x13 e ao fundo o tamanho aumentado de 100x100. . . . .	41
Figura 18 – Arquitetura visual da rede neural convolucional. . . . .	42
Figura 19 – Cristais de $C_3S$ colados por componente esbranquiçado. . . . .	43
Figura 20 – Estrutura dos dados no algoritmo genético. . . . .	44
Figura 21 – Funcionamento da Faster R-CNN . . . . .	45
Figura 22 – Exemplo de máscara e caixa delimitadora . . . . .	46
Figura 23 – Arquitetura da MASK R-CNN. . . . .	46

Figura 24 – Exemplo de um sistema para anotação de máscara, é possível ver um polígono (máscara) que engloba o cristal de $C_3S$ a direita e a esquerda na caixa de texto a sua classe. . . . .	47
Figura 25 – Comparação entre a imagem segmentada manualmente (correta) e a calculada pela Mask R-CNN. . . . .	48
Figura 26 – Cristal de $C_3S$ classificado erroneamente como poro (componente em azul mais a direita) na imagem segmentada pela MASK R-CNN. . . . .	49
Figura 27 – Poro mais acima não foi identificado pela MASK R-CNN. . . . .	49
Figura 28 – Comparação de resultados entre as imagens de segmentação, nesta figura é possível notar que a Mask R-CNN identificou um cristal diferente que estava mais regular com o filtro bilateral. . . . .	51
Figura 29 – Nessa figura é possível notar que na imagem com filtro bilateral um cristal de $C_3S$ não foi classificado erroneamente como poro. . . . .	51
Figura 30 – Nessa figura é possível notar que a componente mais a esquerda da imagem não foi classificada erroneamente como poro. . . . .	52
Figura 31 – Nessa figura é possível notar que o cristal de $C_3S$ mais abaixo e ao lado do poro não foi identificado. . . . .	53
Figura 32 – Nessa figura é possível notar bons resultados e apenas um erro na não identificação de 2 cristais de $C_3S$ ao lado esquerdo do poro. . . . .	54
Figura 33 – Nessa figura é possível notar que o maior cristal de $C_3S$ mais a direita foi identificado porém o cristal a sua esquerda não foi, além disso, a máscara do poro está menor do que deveria. . . . .	55
Figura 34 – Nessa figura é possível notar ótimos resultados. . . . .	56
Figura 35 – Nessa figura é possível notar o cristal de $C_3S$ acima do poro mais a baixo da imagem não identificado. . . . .	56
Figura 36 – Nessa figura houveram erros apenas na separação de cristais de $C_3S$ muito grudados o que pode ser analisado juntamente com um especialista na área como esse cenário seria tratado. . . . .	57



## LISTA DE TABELAS

Tabela 1 – Comparação da monografia com os trabalhos relacionados . . . . .	32
Tabela 2 – Comparação da monografia com os trabalhos relacionados . . . . .	58



## SUMÁRIO

1	<b>INTRODUÇÃO</b>	12
1.1	<b>Clínquer</b>	12
1.2	<b>Indústria 4.0</b>	14
1.3	<b>Automatização da análise de <math>C_3S</math></b>	16
2	<b>OBJETIVOS E RELEVÂNCIA DA PESQUISA</b>	17
2.1	<b>Objetivo geral</b>	17
2.2	<b>Objetivos específicos</b>	17
3	<b>FUNDAMENTAÇÃO TEÓRICA</b>	18
3.1	<b>Indústria 4.0</b>	18
3.2	<b>Complexidade e Necessidade de Automação na Indústria Cimenteira</b>	19
3.3	<b>Sensor Virtual</b>	19
3.4	<b>Microscopia</b>	21
3.4.1	<i>Automação da Microscopia do Clínquer</i>	21
3.5	<b>Aprendizado de Máquinas</b>	22
3.6	<b>Redes Neurais Artificiais</b>	23
3.6.1	<i>Perceptron de uma camada</i>	23
3.6.2	<i>Principais Funções de Ativação</i>	24
3.6.3	<i>Algumas funções de perda</i>	25
3.6.4	<i>Descida do gradiente na estimação dos pesos</i>	26
3.6.5	<i>Perceptron multicamadas</i>	27
3.6.6	<i>Rede Neural Convolutacional</i>	29
3.7	<b>Trabalhos Relacionados à Automação da Microscopia do Clínquer</b>	31
4	<b>FUNDAMENTAÇÃO METODOLÓGICA</b>	33
4.1	<b>Dados para Microscopia do Clínquer</b>	33
4.2	<b>Sistema Autônomo para Microscopia do Clínquer</b>	35
4.3	<b>Exploração das imagens com algoritmos de aprendizado de máquina</b>	36
4.4	<b>Base de dados explorada</b>	36
4.4.1	<i>Pré-processamento geral</i>	37
4.4.2	<i>Segmentação com K-Means</i>	37
4.4.3	<i>Segmentação com K-Nearest-Neighbors( KNN)</i>	38

4.4.4	<i>Identificação dos contornos</i> . . . . .	40
4.5	<b>Classificação com CNN</b> . . . . .	40
4.5.1	<i>Pré-processamento da base de dados de treinamento para a CNN</i> . . . . .	40
4.5.2	<i>Arquitetura da CNN</i> . . . . .	41
4.6	<b>Algoritmo Genético</b> . . . . .	43
4.7	<b>Faster R-CNN</b> . . . . .	44
4.8	<b>Mask R-CNN</b> . . . . .	45
5	<b>RESULTADOS</b> . . . . .	48
5.1	<b>Treinamento da Mask R-CNN</b> . . . . .	48
5.2	<b>Treinamento da Mask R-CNN com Filtro Bilateral</b> . . . . .	49
5.3	<b>Treinamento final da Mask R-CNN</b> . . . . .	52
6	<b>CONCLUSÕES E TRABALHOS FUTUROS</b> . . . . .	59
	<b>REFERÊNCIAS</b> . . . . .	60

## 1 INTRODUÇÃO

O cimento portland é o nome formal para o material comumente usado na construção civil, o cimento. Esse material apresenta-se como um pó fino, que endurece em contato com a água, sendo capaz de se decompor. Na produção de cimento são utilizadas principalmente três tipos de matérias primas, o calcário, a argila e o gesso. Entretanto, dependendo do tipo de cimento, outros materiais podem ser adicionados, tais como escória de alto-forno, pozolanas, materiais carbonáticos etc, pois as matérias primas principais podem não conter todos os elementos necessários para a fabricação de um cimento de boa qualidade. No processo produtivo, calcário, argila e algum tipo de complemento são moídos para dar origem a uma mistura contendo aproximadamente 1 parte de argila para 7 partes de calcário, juntamente com seus complementos, caso sejam necessários. Essa mistura (denominada farinha) é cozida até uma temperatura em torno de 1450°C, obtendo-se um material granulado chamado clínquer. O clínquer é, então, moído com pequena porção de gesso (em torno de 5% do total do produto) sendo o pó fino resultante, de cor cinza, o Cimento Portland. Para obter os diferentes tipos de cimento adicionam-se diferentes tipos de materiais adicionais ao calcário e argila, conforme mencionado anteriormente (GERMAND, 2017).

### 1.1 Clínquer

O clínquer é obtido em um forno rotativo, onde, além do próprio clínquer, são produzidas grandes quantidades de gases de combustão, que são prejudiciais ao meio ambiente, como o  $CO_2$ , por exemplo. Assim, controlar a condição de queima da farinha no interior do forno pode significar, além de um produto final de melhor qualidade, menos emissão de poluentes e menor uso de combustível, que na maioria das vezes são fósseis. Um exemplo de imagem de clínquer pode ser visualizado na Figura 1.

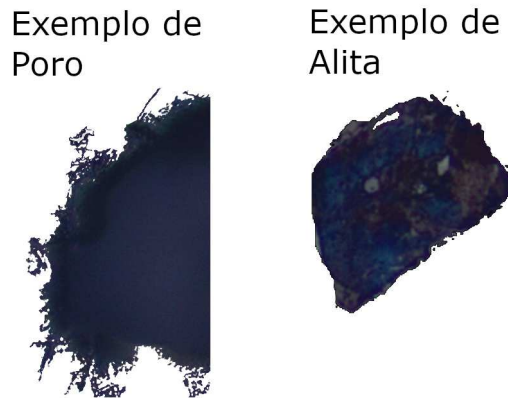
Figura 1 – Exemplo de uma imagem de clínquer



Fonte: Indústria parceira do estudo.

No clínquer existem várias componentes químicas, mas para esse estudo será focado apenas em poros que são componentes químicas que foram queimadas durante o processo de produção e apresentam como principal característica a coloração preta (pode variar levemente dependendo de qual componente química era antes de ser queimada) e a alita ou  $C_3S$  que é um componente composto de  $3CaO$  e  $SiO_2$  e é caracterizado por sua cor azul com partes de cor marrom. Um exemplo de ambas as componentes citadas pode ser visualizado na Figura 2.

Figura 2 – Exemplo de um cristal de um poro e um cristal de alita



Fonte: Autor.

O estudo mineralógico do clínquer já produzido contribui significativamente para a melhora da qualidade do produto final e para a redução do consumo específico de combustível na produção, conseqüentemente, emissão de poluentes (RODRIGUES *et al.*, 2013).

Uma clínquer de boa qualidade necessita atender a vários requisitos estruturais e químicos. Para que essas estruturas sejam analisadas, é executado o processo de análise de sua qualidade, do qual pode ser feito por difração de raios X e/ou microscopia, sendo a última a mais utilizada, provendo mais resultados, além de ser menos custosa (SOUZA *et al.*, 2002). Se o clínquer analisado não é de boa qualidade, é sinal que o sistema que o produziu necessita ser ajustado, pois, um processo produtivo bem controlado torna possível um produto de melhor qualidade com menor custo.

## 1.2 Indústria 4.0

O sistema de produção industrial como um todo tem se tornado altamente complexo, devido à enorme demanda por produtos industriais. Desse modo, a indústria tem sido obrigada a utilizar tecnologia e métodos inteligentes para esses sistemas, visando melhorar a produtividade, minimizando, por exemplo, interrupções na produção. Além disso, é preciso simplificar o processo de supervisão e diminuir os custos de manutenção tanto quanto possível.

Com a introdução dos conceitos de Internet das Coisas (IoT, do inglês *Internet of Things*), sistema físico cibernético (CPS, do inglês *Cyber Physical Systems*), computação em nuvem e inteligência artificial (IA), a automação industrial tem passado por uma grande transformação, possibilitando o advento da indústria 4.0. A indústria 4.0 usa técnicas de computação cognitiva (juntamente com aplicativos IoT) ciência de dados e modelos matemáticos para analisar dados de máquinas, processos e sistemas em tempo real, a fim de monitorar, otimizar e controlar o processo produtivo Masood e Sonntag (2020), Jena *et al.* (2019), Xu *et al.* (2018), Wollschlaeger *et al.* (2017).

Como já mencionado anteriormente, a produção de clínquer envolve as tarefas de combustão de carvão pulverizado, pré-aquecimento e decomposição da matéria-prima. Assim, o desenvolvimento de programas de otimização e controle do processo de pré-decomposição das matérias-primas cimenteiras e a formação dos operadores da produção de cimento são particularmente importantes. Considerando esse cenário, soluções da indústria 4.0 podem trazer grandes benefícios para o processo produtivo do cimento, pois estas soluções podem, por exemplo, melhor gerenciar o uso de energia, evitando desperdícios, lidar com os desafios de alto custo de produção e com toda a complexidade que é inerente à este tipo de indústria, inclusive a alta emissão de  $CO_2$ .

Com o atual nível da tecnologia na área de inteligência artificial, é possível criar um sistema para automatizar a microscopia do clínquer. Isto pode ser realizado utilizando técnicas de visão computacional e algoritmos de aprendizado de máquina para a realização das tarefas de pré-processamento, segmentação e classificação de imagens para caracterização de objetos de interesse (SZUMMER; PICARD, 1998).

O aprendizado de máquina pode ser definido como uma técnica de análise de dados com o propósito de construir modelos analíticos capazes de identificar padrões e, com isso, ter condições de otimizar os acertos em tomadas de decisões. Uma classe de algoritmos mais avançada dentro da área de aprendizado de máquina é formada pelas redes neurais artificiais, que são modelos computacionais que possuem inspiração no funcionamento dos neurônios biológicos. Essas são capazes de aprender e identificar padrões, caso um conjunto de amostras de treinamento seja fornecido de forma apropriada. Existem vários tipos de redes neurais e uma delas é a convolucional (ou CNN *Convolutional Neural Network*) criada por Yann LeCun em 1980 (KALCHBRENNER *et al.*, 2014), sendo essa um tipo de perceptron multicamadas especializado para dados espaciais, como imagens e textos, possuindo como diferencial das



demais redes existentes uma camada de convolução que permite que filtros sejam aplicados de forma a facilitar o aprendizado. Com isso, é possível promover a redução dos dados de entrada e, conseqüentemente, facilitar o seu processamento.

### 1.3 Automatização da análise de $C_3S$

Mesmo com todos esses avanços em algoritmos de aprendizagem, na indústria cimenteira, apesar da grande importância da microscopia para acompanhamento do processo produtivo (HERNANDEZ *et al.*, 2018; MOHAČEK-GROŠEV *et al.*, 2021; BOHÁČ *et al.*, 2021), não foram encontrados na literatura trabalhos que apresentem uma solução automatizada para a realização dessas análises. Deste modo, o objetivo geral desta monografia é segmentar e classificar cristais de  $C_3S$  em imagens microscópicas por meios da construção de uma inteligência artificial que irá ser treinada com imagens pré-processadas, mas com foco na identificação de um componente específico o  $C_3S$ .

Para isso, foi feita uma exploração nas imagens de clínquer utilizando algoritmos de aprendizado como o K-Means (CELEBI, 2014), *K-Nearest-Neighbors* (KNN) (KRAMER, 2013) e Rede Neural Convolutacional (CNN do inglês *Convolutional Neural Network*) (KALCHBRENNER *et al.*, 2014), afim de testar esses algoritmos na segmentação das imagens. Apesar do resultado não ser totalmente insatisfatório, não foi possível separar os cristais de  $C_3S$  nas imagens.

Tal problema foi alcançado de forma satisfatória com a Mask R-CNN (do inglês *Mask Region Convolutional Neural Network*) (HE *et al.*, 2017), que é uma variação da CNN com capacidade de localizar, classificar e segmentar imagens com alta precisão.

Entretanto, por conta da escassez e qualidade das imagens este algoritmo não separou todos os cristais de  $C_3S$  com baixa divisão, ou seja, aglomerado de cristais que possuem divisões discretas entre eles ou quase imperceptíveis. Com mais imagens de qualidade ou algum algoritmo que evidencie essas divisões é possível que a Mask R-CNN consiga separar todos os cristais.

O restante do texto está organizado da seguinte forma. No Capítulo 2 são apresentados, de forma objetiva, os objetivos do trabalho como um todo. No Capítulo 3 é apresentada a fundamentação teórica do problema. Finalmente, no Capítulo 4, são apresentadas as metodologias utilizadas, ferramentas e apresentação do cronograma.

## 2 OBJETIVOS E RELEVÂNCIA DA PESQUISA

### 2.1 Objetivo geral

O objetivo geral é classificar, localizar e aplicar máscara (delimitar precisamente a área das componentes de interesse como  $C_3S$ , poro e etc, um exemplo de máscara será demonstrado no Capítulo 4) em imagens microscópicas através da aplicação de um tipo de rede neural convolucional(CNN do inglês *Convolutional Neural Network*) que é treinada com imagens pré-processadas, além de medir o desempenho geral do modelo gerado assim como a taxa de acerto para cada classe e a área da máscara da qual se trata de cada componente químico a ser identificado, mas com foco na identificação de um componente específico o  $C_3S$ .

### 2.2 Objetivos específicos

- Montar um banco de imagens provenientes de um laboratório de controle de qualidade de uma planta de cimento.
- Segmentar as imagens do banco.
- Aplicar o algoritmo Redes Neurais Artificiais na segmentação de  $C_3S$  nas imagens segmentadas.
- Usar medida de acurácia para avaliar o desempenho do modelo de aprendizado de máquina utilizados.
- Evidenciar características do modelo que podem favorecer ou não a classificação das imagens microscópicas.

### 3 FUNDAMENTAÇÃO TEÓRICA

#### 3.1 Indústria 4.0

O sistema de produção industrial é altamente complexo, devido à enorme demanda por produtos industriais, o que leva a enormes processos produtivos. Deste modo, a indústria foi obrigada a utilizar tecnologia e métodos inteligentes para esses sistemas, visando melhorar e minimizar interrupções na produção. Além disso, é preciso simplificar o processo de supervisão e diminuir os custos de manutenção tanto quanto possível. Essas são medidas importantes, pois também podem levar à satisfação dos clientes, garantir a prevenção dos equipamentos envolvidos e salvar vidas humanas.

Com a introdução dos conceitos de Internet das Coisas (IoT, do inglês *Internet of Things*), sistema físico cibernético (CPS, do inglês *Cyber Physical Systems*), computação em nuvem e inteligência artificial (IA), a automação industrial tem passado por uma grande transformação, possibilitando o advento da indústria 4.0. A indústria 4.0 usa técnicas de computação cognitiva (juntamente com aplicativos IoT) ciência de dados e modelos matemáticos para analisar dados de máquinas, processos e sistemas em tempo real, a fim de monitorar, otimizar e controlar o processo produtivo.

Nesse cenário, a combinação entre os componentes de software e a parte mecânica das máquinas é realizada por meio de CPS. Em geral, os sensores possibilitam a geração de informações a partir das máquinas que chegam aos softwares de forma automática. Desse modo, dados são transferidos automaticamente entre máquinas (físico) e software (ciber). Assim, a análise dos dados produzidos podem possibilitar o monitoramento, controle ou otimização do processo produtivo por meio de uma IA. Esta inteligência poderá, a partir das análises dos dados, gerar *insights*, enviar ou receber comandos ou alertas.

Finalmente, a IoT faz a integração, possibilitando que informações emitidas por um equipamento chegue ao destinatário, o qual poderá dar uma resposta condizente. Para garantir a essa estrutura flexibilidade, agilidade e facilidade na transferência de dados, pode ser usada a computação em nuvem, que é peça fundamental nessa revolução tecnológica. Para maiores detalhes sobre indústria 4.0 e sua implementação, veja, por exemplo, Masood e Sonntag (2020), Jena *et al.* (2019), Xu *et al.* (2018), Wollschlaeger *et al.* (2017).

### 3.2 Complexidade e Necessidade de Automação na Indústria Cimenteira

A pré-decomposição da matéria-prima do cimento (composta principalmente de argila e pedra calcária) é um processo de produção contínuo que apresenta condições complexas e mutáveis na sua realização. Este processo dá origem ao principal componente do cimento, o qual é denominado clínquer. A produção do clínquer envolve a moagem da matéria prima, para dar origem a farinha crua. Essa farinha deve ser aquecida para que seja possível a obtenção de elementos químicos que irão definir a durabilidade e a resistência do cimento. Essa etapa envolve as tarefas de combustão de carvão pulverizado, pré-aquecimento e decomposição da matéria-prima.

Nesse processo, a decomposição do carbonato presente na farinha requer uma temperatura relativamente estável. Se a temperatura estiver muito alta, o pré-aquecedor pode facilmente causar bloqueio. Se a temperatura estiver muito baixa, o forno rotativo (onde ocorre o aquecimento da farinha) ficará sobrecarregado. Além disso, considerando estatísticas mostradas em Maroušek *et al.* (2020), a indústria de cimento está entre os mais significativos emissores de  $CO_2$  do mundo. Isso faz com que o controle do processo produtivo, afim de evitar aquecimento excessivo dentro do forno, seja uma questão de sustentabilidade que envolve todo o planeta. Detalhes sobre a produção do clínquer podem ser vistos, também, em Ma *et al.* (2020), Sutawijaya *et al.* (2021), Pieper *et al.* (2021).

Assim, o desenvolvimento de programas de otimização e controle do processo de pré-decomposição das matérias-primas cimenteiras e a formação dos operadores da produção de cimento são particularmente importantes. Considerando esse cenário, soluções da indústria 4.0 podem trazer grandes benefícios para o processo produtivo do cimento. Pois essas soluções podem, por exemplo, melhor gerenciar o uso de energia, evitando desperdícios, lidar com os desafios de alto custo de produção e com toda a complexidade que é inerente à este tipo de indústria, inclusive a alta emissão de  $CO_2$ .

### 3.3 Sensor Virtual

Como descrito na Seção 3.1, para que seja possível automatizar um processo produtivo, faz-se necessária a obtenção de dados a partir das máquinas envolvidas no processo de produção. Muitas vezes, é possível a obtenção desses dados via sensores. Na produção de cimento, muitas variáveis envolvidas podem ser aferidas por sensores, como a quantidade de

matéria-prima utilizada (que pode ser pesada antes de ser utilizada na produção), a temperatura atual de entrada e saída do forno, potências de motores de ventiladores utilizados para movimentar o produto dentro do forno, e muitas outras. A partir dessas variáveis, modelos de aprendizado de máquinas podem ser utilizados para fornecerem *insights* sobre o processo. Um exemplo pode ser visto no trabalho de Shi *et al.* (2021), onde é feita uma abordagem com redes neurais convolucionais para fazer previsão síncrona do consumo de eletricidade e carvão na produção de cimento. No entanto, existem variáveis que não podem ser aferidas via sensor convencional.

Na indústria cimenteira, parâmetros de qualidade por exemplo, em geral, são aferidos em laboratórios, sendo necessário um tempo determinado para sua realização. Desse modo, caso essas variáveis precisem ser utilizadas como resposta para fornecerem evidências sobre o estado do processo produtivo, a solução automática fica prejudicada. Para amenizar esse problema, são propostos modelos de aprendizado de máquinas para inferirem sobre essas variáveis, usando, para isso, outras variáveis que possam estar correlacionadas com essa resposta. Esses procedimentos de estimação são chamados de *soft sensor* ou sensor virtual. Um exemplo de sensor virtual pode ser visto em Xu *et al.* (2021), onde os autores propõem um sensor virtual para estimar a zona de queima dentro do forno rotativo, uma vez que essa variável não pode ser aferida diretamente, devido as altas temperaturas que podem ocorrer dentro do forno.

A necessidade de monitoramento da zona de queima é devida ao fato de que essa variável afeta o status operacional do forno, a qualidade do clínquer, o consumo de energia e as emissões de poluição. A abordagem descrita no trabalho dos autores pode ser vista como um trabalho relacionado a esta monografia, pois uma forma de obter algumas informações sobre as temperaturas que foram alcançadas dentro do forno rotativo na produção de um determinado clínquer é utilizando a microscopia, a qual é descrita nas seções seguintes.

A microscopia do clínquer, Além de fornecer informações sobre a condição de queima dentro do forno, também possibilita a obtenção de informações sobre a farinha utilizada na produção, ou seja, condições de moagem da matéria-prima. Adicionalmente, é possível tirar conclusões sobre a qualidade do cimento resultante do processo de moagem do clínquer com adição de gesso. Todas estas informações são de difíceis, ou até mesmo impossíveis, de serem obtidos de forma direta. No entanto, a realização da microscopia é uma tarefa totalmente manual, necessitando de automação. Uma vez automatizada, a técnica pode ser vista como um sensor virtual, pois fará uso de características fornecidas por imagens microscópicas para estimação de certas condições de operação do forno e processos de moagem que podem acarretar problemas

como a má qualidade do produto final, a emissão excessiva de  $CO_2$ , além de desperdícios de matéria-prima, água e energia.

### 3.4 Microscopia

O microscópio é um importante instrumento, pois muitas vezes auxilia e aperfeiçoam a visão humana, sendo a peça fundamental na microscopia. Atende-se por microscopia a análise de amostras de certos elementos com o uso de um microscópio. Os avanços na tecnologia e digitalização deram início a novas maneiras de aprimorar a pesquisa por meio de microscopia digital e análise de imagens. Microscópios modernos permitem a digitalização das imagens, que podem ser armazenadas em repositórios virtuais para serem visualizadas em computadores em vez de microscópios.

Compartilhamento mais fácil e rápido de imagens, melhor armazenamento e preservação da qualidade das imagens e anotação de características dos espécimes analisados são apenas algumas das vantagens dessa tecnologia. Combinados com o desenvolvimento de software para análise digital, as imagens digitais abrem ainda mais o caminho para o desenvolvimento de ferramentas que extraem dados quantitativos de estudos baseados em imagens. Esse tipo de estudo vem sendo desenvolvido em muitas áreas, como estudos de patologias humanas (KRISHNAMURTHY *et al.*, 2019), em arqueologia (MARTÍN-VIVEROS; OLLÉ, 2020), análise de qualidade de frutas (GAO *et al.*, 2021), e etc.

Em análise de minérios de ferro, Iglesias *et al.* (2018) apresentam um sistema automático para caracterização mineralógica com base em microscopia digital. Após a obtenção automática de imagens utilizando um microscópio de luz refletida motorizado, os autores usam classificadores para discriminar importantes componentes desse minério, emitindo de forma automática relatórios contendo informações quantitativas das fases mineral e textual. Na indústria, trabalhos como o de Kim *et al.* (2020) mostram aplicações da microscopia digital para controle de qualidade como sendo uma solução promissora e poderosa.

#### 3.4.1 Automação da Microscopia do Clínquer

Na indústria cimenteira, apesar da grande importância da microscopia para acompanhamento do processo produtivo (HERNANDEZ *et al.*, 2018; MOHAČEK-GROŠEV *et al.*, 2021; BOHÁČ *et al.*, 2021), não foram encontrados na literatura trabalhos que apresentem uma

solução automatizada para a realização dessas análises.

Com o atual nível da tecnologia na área de inteligência artificial, é possível criar um sistema para automatizar a microscopia do clínquer. Isto pode ser realizado utilizando técnicas de visão computacional e algoritmos de aprendizado de máquina para a realização das tarefas de pré-processamento, segmentação e classificação de imagens para caracterização de objetos de interesse (SZUMMER; PICARD, 1998).

A automação pode ser composta por um conjunto de algoritmos para executar o pré-processamento e a segmentação da imagem analisada. Um exemplo de método para aprendizado de máquina, que faz parte do grupo dos algoritmos de agrupamento, é denominado K-Means, que é um método de agrupamento não supervisionado. Esse método agrupa os dados com base na sua similaridade e distâncias de um ponto central que irá representar o centro de um determinado grupo. Para identificação dos elementos destacados na etapa de segmentação, pode ser utilizado um classificador (ZHENG *et al.*, 2018), ou seja, um método supervisionado.

O foco principal deste trabalho é a aplicação de uma rede neural convolucional para realização da classificação da Alita a fim de contribuir com a automação da microscopia do clínquer.

### **3.5 Aprendizado de Máquinas**

Com a refinação de métodos estatísticos que se deu em meados de 1950, ocorreu a pesquisa pioneira em aprendizado de máquina, conduzida usando algoritmos simples. Porém, o termo aprendizado de máquina foi cunhado em 1959 por Arthur Samuel, um funcionário da IBM (*International Business Machines Corporation*), o qual foi pioneiro no campo de jogos de computador e inteligência artificial. Então, em meados de 1960, os métodos Bayesianos (GELMAN *et al.*, 1995) para inferência a partir de modelagem probabilística são introduzidos em aprendizado de máquina, surgindo, assim, os primeiros algoritmos dessa área. Em 1950 foi desenvolvida a ideia do neurônio artificial por Warren McCulloch e Walter Pitts, componente base das redes neurais, que surgiram por volta de 1951, com a criação da primeira rede neural. Esse algoritmo foi desenvolvido por Marvin Minsky and Dean Edmonds em (MINKSY; EDMONDS, 1954), tendo sido denominada "SNARC".

Com o avanço das pesquisas científicas na área de aprendizado de máquinas, mais algoritmos foram surgindo como, por exemplo, o SVM (CORTES; VAPNIK, 1995) na década iniciada em 1990, o qual foi desenvolvido pelo laboratório "AT&T Bell". O aprendizado

de máquina pode ser definido como uma técnica de análise de dados com o propósito de construir modelos analíticos capazes de identificar padrões e, com isso, ter condições de otimizar os acertos em tomadas de decisões. Os algoritmos de aprendizado podem ser divididos em supervisionados e não supervisionados, onde os supervisionados são aqueles que requerem que a classe a ser identificada ou estimada seja informada ao algoritmo em uma etapa denominada "treinamento". Enquanto os não supervisionados são aqueles que não requerem essa informação, sendo compostos pelos algoritmos de agrupamento com base em associação por algum tipo de similaridade.

### 3.6 Redes Neurais Artificiais

As redes neurais artificiais são modelos computacionais que funcionam de modo similar ao cérebro humano. Uma rede neural é representada visualmente como um grafo, ou seja, um conjunto de vértices (V) que representam os neurônios e as arestas (E) onde essas fazem as ligações entre os vértices, sendo responsáveis pelo fluxo de dados em toda a rede. As redes neurais artificiais são capazes de aprender e identificar padrões, caso amostras de treinamento sejam fornecidas de forma apropriada. Os neurônios são peças fundamentais para a operação de uma rede neural. Esta seção fornece uma descrição do modelo neuronal, onde são apontados seus elementos básicos e algumas funções de ativação de um dado neurônio.

#### 3.6.1 *Perceptron de uma camada*

A *perceptron* de uma camada é o tipo mais simples de rede neural, composta por apenas um neurônio artificial, o qual foi criado por Frank Rosenblatt nas décadas de 1950 e 1960. Essa rede possui um número  $n > 0$  de entradas que produzem uma única saída. A Figura 3 mostra uma rede *perceptron* composta por um único neurônio artificial (ou *perceptron*), onde pode ser visto que seu funcionamento se inicia nas entradas  $(x_1, \dots, x_n)$ , com cada uma delas possuindo, respectivamente, um peso  $w_i$ , para  $i = 1, \dots, n$ . Esses pesos definem, neste caso, a contribuição de cada entrada na resposta fornecida, ou saída da rede. No centro dessa rede é efetuada a soma de todas as entradas ponderada pelos pesos  $(\sum_1^n x_i * w_i)$ . Assim, nessa soma,  $x_i$  representa a  $i$ -ésima entrada da rede e  $w_i$  o seu peso associado. Por fim, uma função de ativação é utilizada para obtenção da saída. No exemplo visto na Figura 3, a função de ativação utilizada irá retornar 0 ou 1 caso o resultado da soma supere o limiar pré-definido na função de ativação,



como por exemplo:

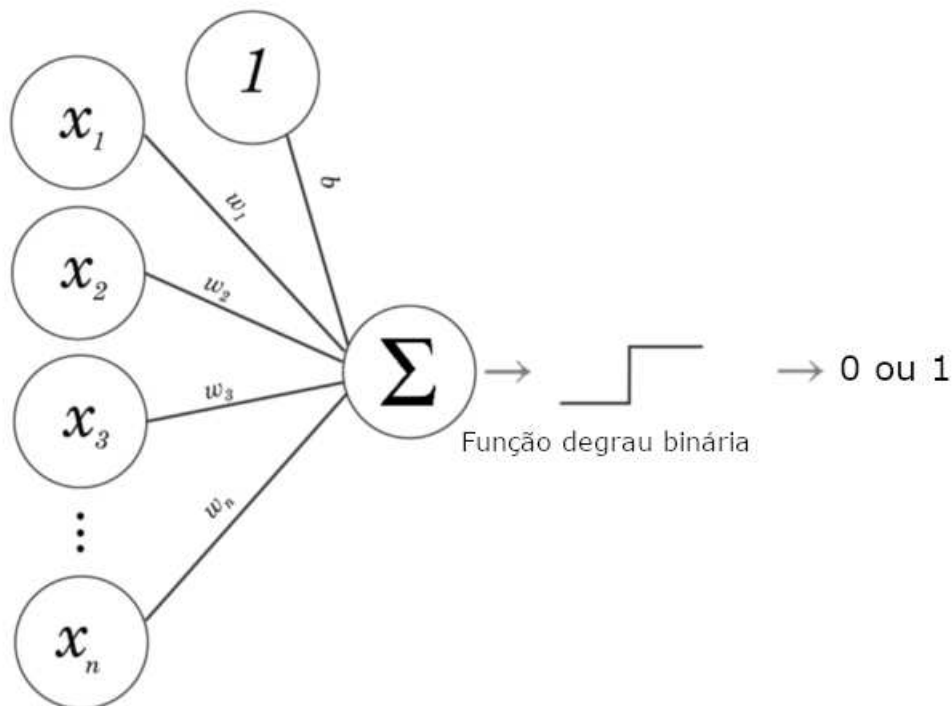
$$\begin{aligned} 0 & \text{ se } \sum_1^n x_i * w_i \leq \text{limiar} \\ 1 & \text{ se } \sum_1^n x_i * w_i > \text{limiar}. \end{aligned} \quad (3.1)$$

A equação anterior não leva em consideração um importante parâmetro para uma rede neural. Ainda na Figura 3, pode ser observado um nó com valor 1, que proporciona a entrada do viés, denotado por  $b$ , que tem como objetivo controlar a entrada líquida da função de ativação, promovendo uma regularização dessa função. Com isto, tem-se a equação:

$$\begin{aligned} 0 & \text{ se } b + \sum_1^n x_i * w_i \leq \text{limiar} \\ 1 & \text{ se } b + \sum_1^n x_i * w_i > \text{limiar}, \end{aligned} \quad (3.2)$$

o viés irá modificar a saída conforme seu sinal seja negativo ou positivo.

Figura 3 – Exemplo de um perceptron de uma camada.



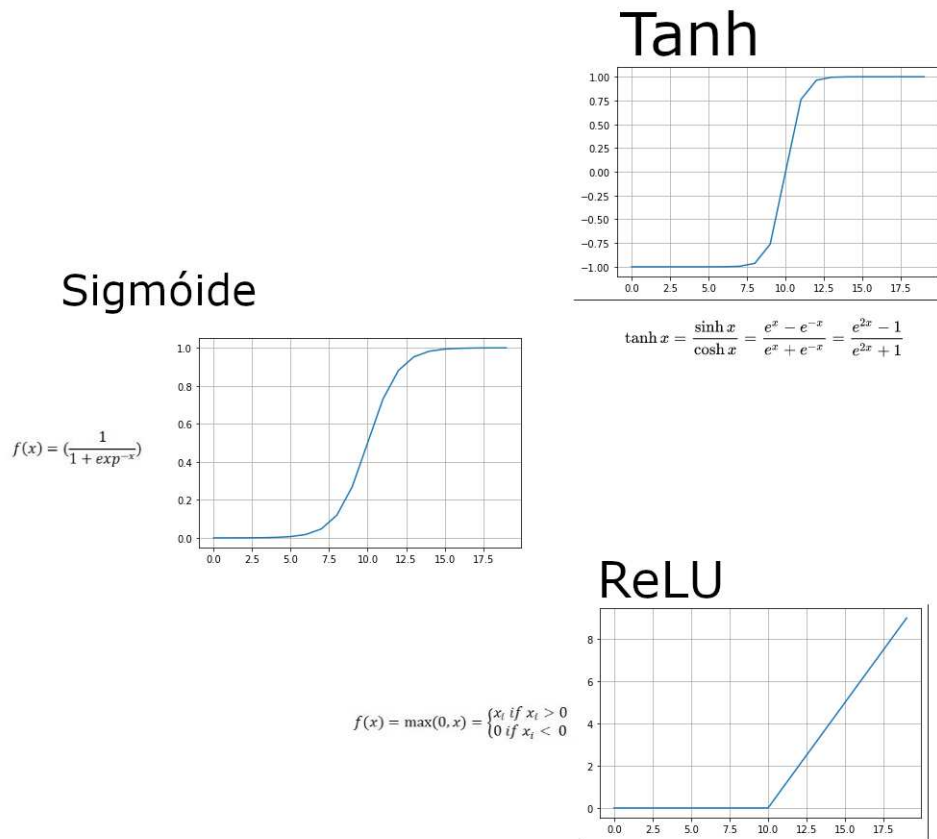
Fonte: Dhalla (2021)

### 3.6.2 Principais Funções de Ativação

A função de ativação retorna a saída de um neurônio artificial, definindo se esse neurônio será ativado ou não. Em outras palavras, essa função é usada para decidir se a

informação que o neurônio recebe é relevante ou deve ser ignorada. Alguns exemplos de funções de ativação podem ser vistos na Figura 4, onde também pode ser visto um gráfico representando seu comportamento.

Figura 4 – Principais funções de ativação e seus respectivos comportamentos em gráfico.



Fonte: Autor.

### 3.6.3 Algumas funções de perda

A função de perda (ou função de erro) (ROSASCO *et al.*, 2004) pode ser utilizada para calcular o gradiente da rede e atualizar os pesos minimizando o erro. O gradiente é o vetor que indica a direção em que o erro aumenta mais rapidamente, assim como a direção oposta a qual esse erro diminui. Essa função sempre é executada no fim do fluxo da rede, após a função de ativação da saída da rede. As principais funções de erro utilizadas na prática são listadas a seguir.

- Função Erro Quadrático Médio (MSE, do inglês *mean squared error*), dada por:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2, \quad (3.3)$$

em que  $y_i$  é o valor predito e  $y'_i$  o valor observado para uma determinada classe (ou grupo).

- Função Entropia Cruzada Binária (traduzido do inglês *binary crossentropy*)

$$\text{BCE} = -\frac{1}{n} \sum_{i=1}^n (y_i * \log(p(y_i)) + (1 - y_i) * \log(1 - p(y_i))), \quad (3.4)$$

em que  $\log(p(y_i))$  é o logaritmo da probabilidade de ocorrência de saída 1 e  $\log(1 - p(y_i))$  o logaritmo da probabilidade da ocorrência 0, na função binária.

- Entropia Cruzada Binária para problema multi-classe

$$\text{BCEM} = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m y_{ij} \log(p(y_{ij})), \quad (3.5)$$

esta função é a modificação da BCE para tratamento de problemas multi-classe, ou seja, onde ocorrem mais de dois tipos de classificação possível para cada objeto. Nessa equação, é considerada para cada classe um vetor contendo  $m$  componentes com valor 1 na posição que representa a classe e zero nas demais componentes do vetor. Considerando  $n$  objetos a serem classificados, tem-se uma matriz  $n \times m$ .

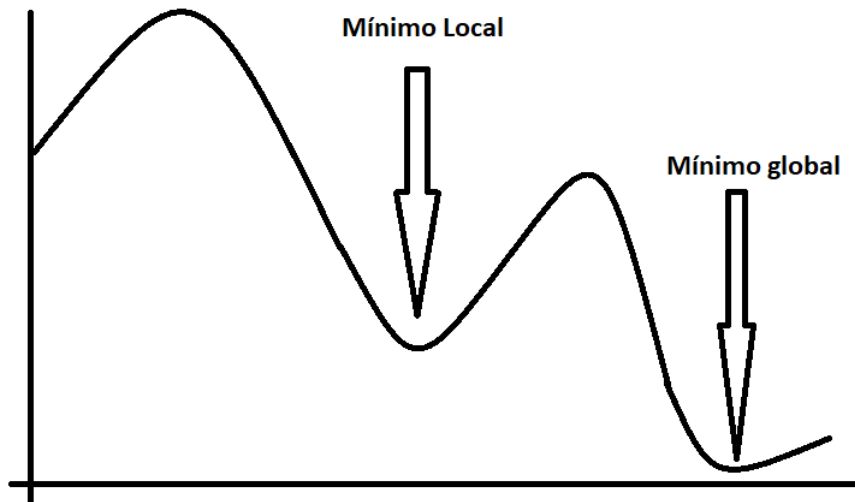
### 3.6.4 Descida do gradiente na estimação dos pesos

A descida do Gradiente proposta por **Augustin-Louis Cauchy** em 1847 é um algoritmo para otimizar soluções de problemas utilizando funções diferenciais iterativamente. Dado um gradiente calculado a partir da função de erro, são realizados passos na direção oposta do gradiente, até que seja atingido o ponto de mínimo. Entretanto, o algoritmo pode encontrar dois tipos de mínimo, o local e o global, como ilustrado na Figura 5. O ponto de mínimo local é aquele que assume o menor valor em uma vizinhança de pontos, enquanto o mínimo global é o menor ponto possível de todo os pontos contidos no gradiente. Os principais otimizadores para este tipo de problema são listados a seguir.

- Descida do Gradiente em Lote (conhecido como *Batch Gradient Descent*, do inglês), este método usa a média dos gradientes de todos os registros na base de dados para atualizar os pesos na rede.
- Descida do Gradiente Estocástico (SGD) que tem como estratégia calcular a direção do gradiente para cada elemento da base de dados, forçando, conseqüentemente, a identificação de mais pontos mínimos em um menor espaço de tempo. Dessa forma, uma quantidade maior de atualização de pesos é realizada no mesmo espaço de tempo.
- Descida do Gradiente por Mini Lote (conhecido como *Mini Batch Gradient Descent*, do inglês), é uma mistura do SGD e do método da Descida do Gradiente em Lote. Utiliza um

conjunto com número fixo de exemplos da base de dados de treinamento (mini lote) para obtenção do gradiente, possibilitando a obtenção do gradiente médio de cada mini lote resultante para, então, atualizar os pesos.

Figura 5 – Exemplo de pontos de máximo e mínimo que podem ser encontrados na descida do gradiente.

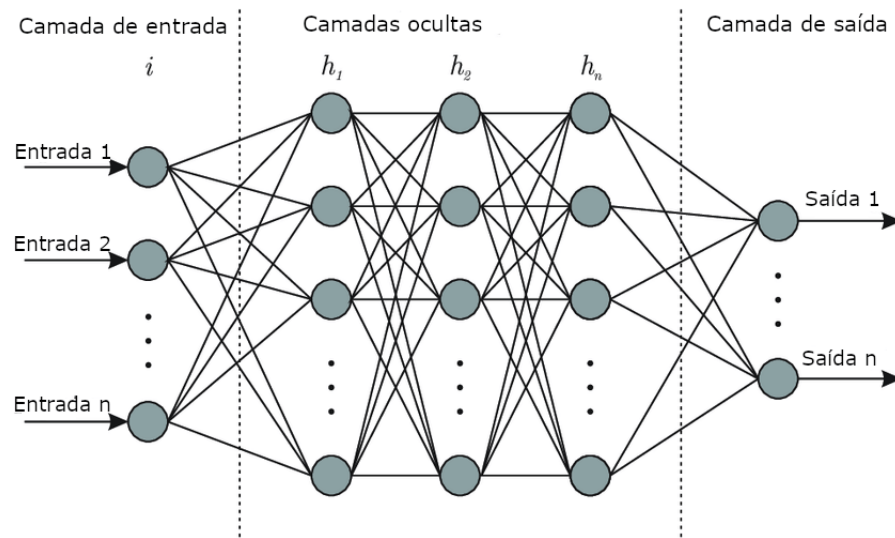


Fonte: Autor.

### 3.6.5 *Perceptron multicamadas*

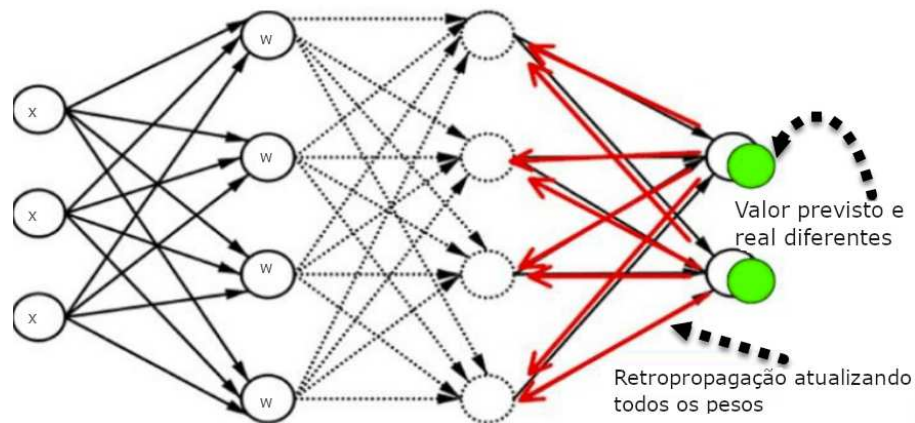
O perceptron multicamadas consiste em um perceptron com duas ou mais camadas. Uma camada pode ser definida como uma coleção de neurônios que operam em conjunto em determinada profundidade na rede, sendo as camadas no interior da rede, entre a primeira e a última, sendo denominadas “camadas ocultas”, como ilustrado na Figura 6. As camadas conferem a rede à capacidade de identificar tendências não lineares no conjunto de dados. Em problemas de classificação, as camadas possibilitam a identificação de grupos (ou classes) que não podem ser separados considerando uma única reta.

Figura 6 – Exemplo de uma rede neural multicamadas com três camadas ocultas.



Fonte: Bre (2017)

Figura 7 – Exemplo do fluxo de informações da retropropagação no interior da rede. Começando dos neurônios de saída até os de entrada.



Fonte: Johnson (2022)

Uma rede neural multicamadas pode ter uma etapa adicional em sua estrutura que é a atualização dos pesos, um exemplo de algoritmo para a atualização dos pesos é a retropropagação do erro, a qual é composta por duas outras etapas:

- a do processamento direto, em que uma entrada é alimentada para a rede neural da qual é processada, fluindo pela rede de camada em camada não alterando os pesos e
- a do processamento reverso, em que um sinal de erro é calculado assim que a saída da

rede é gerada, sendo em seguida propagado de volta, da camada de saída até a camada de entrada, ajustando os pesos de acordo com uma regra de correção de erro,

A retropropagação do erro ocorre após o cálculo do gradiente e a execução da sua descida, sendo o resultado desse algoritmo usado na atualização dos pesos. A Figura 7 apresenta uma ilustração da retro-propagação, onde pode ser visto o fluxo de informações na estrutura da rede.

### 3.6.6 Rede Neural Convolutacional

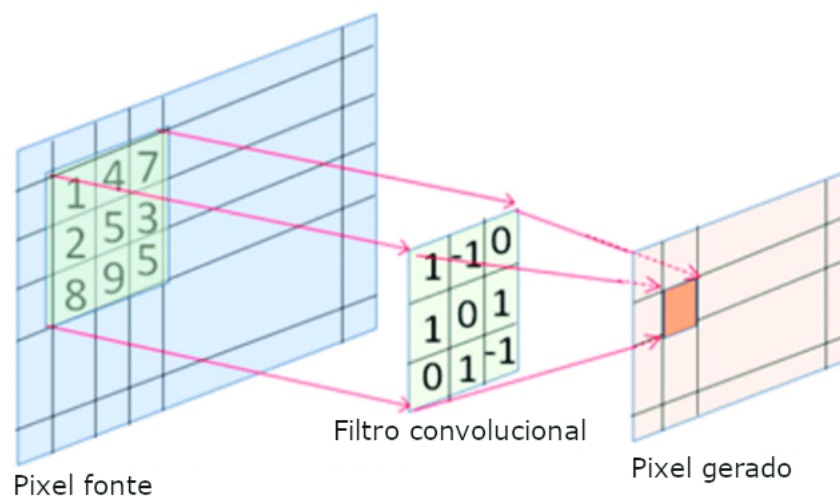
A rede neural convolutacional (ou CNN) criada por **Yann LeCun** em 1980 (KALCHBRENNER *et al.*, 2014), é um tipo de perceptron multicamadas especializado para dados espaciais, como imagens e textos. A principal diferença entre a CNN e um perceptron multicamadas, descrito na seção anterior, é a camada de convolução que permite que filtros sejam aplicados de forma a facilitar o aprendizado, pois com isso é possível promover a redução dos dados de entrada e, conseqüentemente, facilitar o seu processamento.

Considerando um sistema de cores como RGB (do inglês, *Red, Blue e Green*), os dados são originalmente armazenados em três matrizes  $L \times K$ , em que  $L$  representa o número de pixels nas colunas e  $K$  o número de pixels nas linhas. Desta forma, tem-se uma imagem com  $L \times K$  pixels. Cada matriz representa uma dimensão de cor, e a combinação das três dá origem as cores que são exibidas na imagem. Muitas vezes, este formato não é o mais conveniente na análise dos dados da imagem. Assim, cada matriz pode ser convertida para um vetor. Com isso pode ser pensado um conjunto de dados composto por três linhas exibindo os dados de dimensão de cor e outras duas linhas contendo as coordenadas dos pixels, tudo em uma única matriz ou dataframe.

Em uma CNN para análise e classificação de imagens, nas camadas de convolução, pixels de entrada são conectados a uma camada de neurônios ocultos, de forma que regiões pequenas de pixels sejam conectadas a diferentes neurônios. Considerando a forma original dos dados de imagens (matrizes contendo as dimensões de cores), pode ser pensada uma situação como ilustrada na Figura 8. Nasa figura pode ser percebida uma janela de pixels na imagem de entrada. Essa situação ilustra o que é chamado de campo receptivo local, o qual é deslizado, percorrendo toda a imagem, sendo que para cada campo receptivo local diferente deve existir um neurônio oculto diferente na primeira camada oculta. Observe que, se a janela construída contém  $10 \times 10$  pixels, ocorrerá uma redução dos pixels de  $L \times K$  para  $(L - 9) \times (K - 9)$ , caso seja considerado o deslizamento por um pixel de cada vez. Este deslizamento pode ocorrer

considerando um número maior de pixels, como 2, 3, etc. O número de pixel para o deslizamento pode ser considerado como um hiperparâmetro conhecido como *stride length*, uma estratégia para estimação deste hiperparâmetro pode ser vista em Nozaki e Watanabe (2020). O tamanho do campo local também pode ser visto como um hiperparâmetro, em geral, campos locais maiores podem trazer melhor desempenho para imagens com quantidade grande de pixels. Estudos sobre a escolha do campo local podem ser vistos em Keshishian *et al.* (2020), Luo *et al.* (2016).

Figura 8 – Ilustração da aplicação de um filtro convolucional.



Fonte: Commons (2019)

Após a extração de características necessárias ocorrida em uma ou mais camadas de convolução, todas as características obtidas na última camada de convolução são transformadas em um vetor unidimensional, para que seja possível a realização da conexão total. Assim, a camada de saída pode realizar a classificação da imagem de entrada para uma das classes, ou grupos, existentes. Em geral, uma CNN básica consiste de uma camada convolucional, uma camada de *pooling* e uma camada de conexão completa.

Técnicas de convolução e *pooling*, em uma CNN, são responsáveis por filtrar informações espectrais captadas por sensores que são, geralmente, sujeitas a vários distúrbios. Esses distúrbios, caso não sejam corrigidos, podem levar à classificação de objetos de mesma classe para classes distintas e vice-versa. Para isso, pode ser usado um operador denominado *pooling* médio global ou máximo global. Xie *et al.* (2021), além de dar uma descrição matemática do funcionamento da camada com *pooling*, propõe, além da camada com *pooling* de forma usual para CNN, uma última camada de *pooling* usando superpixel. Os autores afirmam obter

bons resultados em classificação de objetos via imagem devido à alta probabilidade de pixels pertencentes ao mesmo superpixel pertencerem à mesma classe, além de promoverem uma maior obtenção de informações espaciais. A estrutura da rede utilizada para alcançar o objetivo deste trabalho é proposta levando em consideração as especificidades das imagens analisadas, estando descrita no Capítulo 4.

### 3.7 Trabalhos Relacionados à Automação da Microscopia do Clínquer

Nesta seção é apresentada uma revisão de literatura, englobando trabalhos que abordam a classificação de imagens de clínquer via algoritmos de aprendizado de máquina. Também é disponibilizada uma tabela para nível de comparação entre os trabalhos encontrados na literatura e a abordagem desta monografia, afim de destacar o diferencial da pesquisa.

Stutzman *et al.* (2016) apresenta um estudo sobre a segmentação e classificação de imagens de clínquer, com intuito de realizar análises quantitativa (contagem de componentes químicos presentes no clínquer). Os autores propõem o uso do algoritmo *Statistical Region Merging* (SRM) (NOCK; NIELSEN, 2004) para a etapa de segmentação da imagem. Para a realização do trabalho, foi utilizada uma grande base de dados de imagens obtidas a partir de um microscópio de varredura (STUTZMAN *et al.*, 1999), o qual é capaz de capturar imagens microscópicas de alta resolução. No entanto, este tipo de microscópio é de difícil manuseio e custos elevados. Como resultado, o algoritmo alcançou uma alta precisão na classificação, chegando a 97%. Entretanto, esse valor foi baseado em apenas uma imagem o que o próprio autor afirma que o valor não representa a precisão geral do algoritmo.

Jourlin *et al.* (2001) apresenta um estudo sobre a segmentação de imagens de clínquer para identificar os componentes químicos presentes nas mesmas. Os autores fizeram uso do algoritmo *Logarithmic image processing* (LIP) (JOURLIN; PINOLI, 2001), utilizando uma base de dados com imagens em tons de cinza. A técnica mostrou bons resultados, porém o autor informa que o método não é totalmente automatizado, devido à reprodução limitada da corrosão química usada para evidenciar as fases mineralógicas presentes no clínquer analisado. Além disso, as variações nas características (forma, matiz, tamanho) dos elementos a serem segmentados não permitem tal automação. O método apresentado pelos autores necessitaria de adaptações para avaliar clínquer de origens diferentes para que seja possível um processo de análise automatizado.

Além desses trabalhos, não foram encontrados na literatura outras pesquisas que



tratassem da automação da microscopia do clínquer. Assim, na Tabela 1 mostra os aspectos que diferenciam os trabalhos que foram encontrados da abordagem realizada nesta pesquisa. Os aspectos marcados com X indicam sua existência nos trabalhos e o X destacado de vermelho simboliza o diferencial desta monografia em relação aos trabalhos relacionados.

Tabela 1 – Comparação da monografia com os trabalhos relacionados

<b>Lista de atividades</b>	<b>Stutzman et al. (2019)</b>	<b>Jourlin et al. (2020)</b>	<b>TCC Renê Michel</b>
<b>Pré-processamento das imagens</b>	x	x	x
<b>Divisão da base de dados em treino e teste</b>			x
<b>Necessita adaptação</b>		x	
<b>Classificação automatizada dos componentes</b>			x

Fonte: elaborado pelo autor (2022).

## 4 FUNDAMENTAÇÃO METODOLÓGICA

Esta seção é dedicada à apresentação do cenário de aplicação dos algoritmos. Aqui é descrito o processo que dá origem as imagens analisadas, equipamento utilizado, técnicas de pré-processamento, segmentação e classificação.

### 4.1 Dados para Microscopia do Clínquer

Para que seja possível a realização da microscopia, porções de clínquer são coletadas e misturadas a uma resina de elevado índice de refração. Essas amostras são prensadas e distribuídas em pequenos espaços de um mesmo molde, como ilustrado na Figura 9. Uma face, a que fica descoberta, é deixada propositalmente sem proteção (lamínula, ilustrada na Figura 9), para que possam ser aplicadas técnicas de polimentos e ataques químicos. O uso de alguns produtos químicos é, muitas vezes, necessário para que os elementos presentes na imagem sejam evidenciados. As lamínula produzidas são levadas para serem analisadas em microscópio. Esse processo pode possibilitar a observação do estado mineralógico e químico do clínquer, com isso pode ser possível extrair informações importantes a respeito de sua produção. Essas informações podem ser usadas para ajustar o processo de produção, caso seja percebida alguma característica que indique problemas no processo produtivo (SOUZA *et al.*, 2002).

Figura 9 – Ilustração do molde exibindo duas lamínulas construídas para obtenção de imagens para a realização da microscopia.



Fonte: Autor (2022).

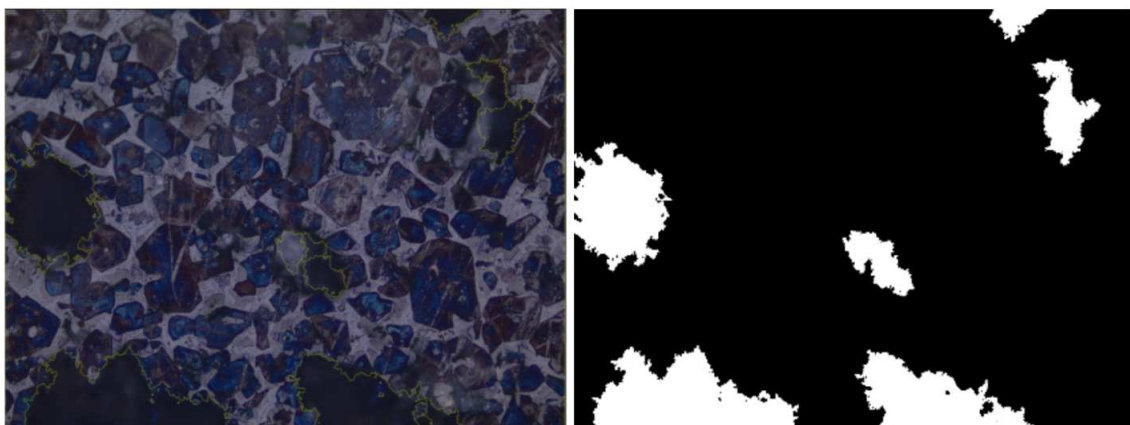
O método proposto neste trabalho se inicia após a captura da imagem microscópica a partir do clínquer disposto em lamínulas. Com isso, deseja-se obter um método que possa contribuir com a automação da criação de relatórios com importantes informações fornecidas pela microscopia. Como exemplo, a Figura 10 mostra do lado direito uma imagem microscópica

obtida a partir da lamínula, após esta ser exposta a amônia por 35 segundos a fim de que os cristais possam obter a coloração azul. Nesta imagem, localizada a direita da figura, é possível perceber elementos com as colorações azul e marrom, esses elementos são os cristais de alita (ou  $C_3S$ ). A coloração azul diz respeito a reatividade do clínquer, que é uma característica a ser levada em consideração na microscopia. Adicionalmente, podem ser percebidas regiões pretas, destacadas de amarelo, que indicam a presença de poros (que traz informações sobre a porosidade do material). Os poros foram identificados na imagem de forma automática, com o uso dos métodos *canny* e *watershed* no processo de segmentação, veja detalhes desses métodos em (LI *et al.*, 2003; NGUYEN *et al.*, 2003; CANNY, 1986).

A aplicação dos métodos *canny* e *watershed* permite a segmentação dos poros, como ilustrado na imagem da direita da Figura 10. Esta segmentação permite a análise da porosidade de forma automática. No entanto, a segmentação dos cristais de alita é uma tarefa mais complexa, devido aos cristais possuírem mais de uma cor, além de muitas vezes aparecerem unidos.

Além do estudo da distribuição das diferentes formas dos cristais de alita, é sabido que outras análises devem ser realizadas, a fim de se obter uma automação completa da microscopia. Tendo em vista que o estudo desse cristal aparenta ser o mais desafiador, sendo também a parte mais importante da microscopia, este trabalho está focado, principalmente, na parte de obter a base que possibilita a análise da forma e distribuição dos cristais de alita.

Figura 10 – Ilustração da imagem microscópica após ataque com amônia por 35 segundos.



Fonte: Autor e empresa parceira da pesquisa.

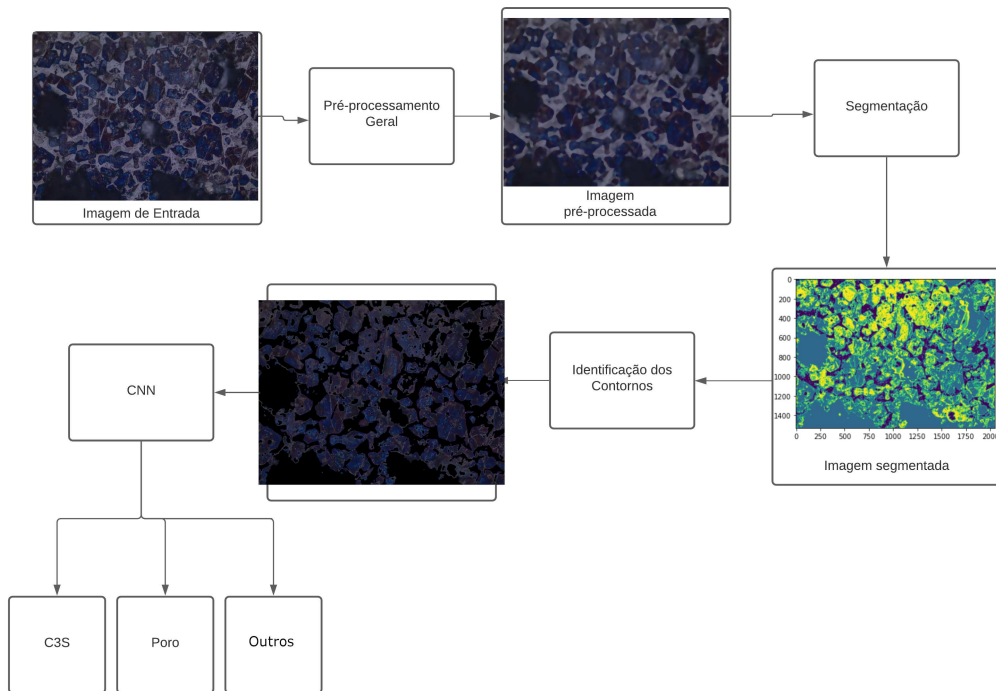
Atualmente, está sendo estudado um banco de dados contendo trezentas imagens como a da esquerda da Figura 10, que foram obtidas a partir de um lote de amostras, ou seja um

conjunto de amostras coletadas ao mesmo tempo. Todas as imagens contêm informações sobre a condição de captura, como lente, zoom e iluminação do aparelho utilizado na captura. As imagens foram obtidas por um profissional capacitado para realização da microscopia. Para este trabalho, mais imagens serão obtidas a partir de outras amostras, a fim de capturar a variabilidade promovida pela condição de operação do forno e moinho de farinha da planta cimenteira. Todas as imagens possuem o mesmo zoom, diferindo apenas na qualidade das componentes presentes nas imagens, contraste e a presença de borões.

## **4.2 Sistema Autônomo para Microscopia do Clínquer**

O sistema autônomo para a Microscopia do Clínquer deve conter técnicas de visão computacional que fazem uso de algoritmos de aprendizado de máquina. Estes algoritmos possuem, basicamente, cinco fases: o pré-processamento geral, a segmentação, a identificação dos contornos de elementos investigados, classificação e análises de tais elementos. As análises podem envolver mensurações de diâmetros, quantidades e distribuição dos cristais, entre outras abordagens. O foco principal deste trabalho consiste nas etapas desde o pré-processamento da imagem até sua classificação, no que diz respeito aos cristais de  $C_3S$ . Estas etapas estão resumidas no diagrama apresentado na Figura 11, sendo melhores descritas nas seções seguintes.

Figura 11 – Diagrama ilustrativo para as etapas de processamento realizadas nas imagens de clínquer até a fase de classificação dos componentes mineralógicos.



Fonte: Autor (2022).

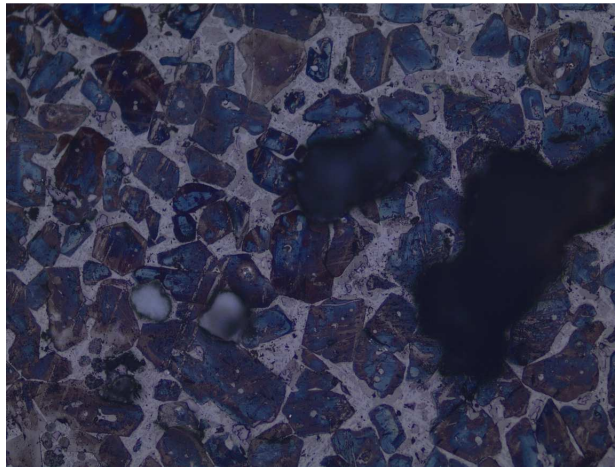
### 4.3 Exploração das imagens com algoritmos de aprendizado de máquina

Esta seção é destinada a exploração da base de dados por algoritmos de agrupamento e classificação. Esta exploração é uma etapa importante da pesquisa, uma vez que não existem na literatura algoritmos treinados para segmentar e classificar cristais de  $C_3S$ .

### 4.4 Base de dados explorada

A base de dados exploradas aqui são imagens microscópicas de clínquer com alta presença de  $C_3S$ , cujo método de obtenção é descrito na seção anterior. As imagens foram coletadas diretamente de uma planta cimenteira e são geradas por um microscópio digital. O formato das imagens na hora da captura é TIF (do inglês *Tag Image File*), o qual mantém a qualidade da imagem. Assim, essas são imagens pesadas e ricas em detalhes e comumente medindo em torno de 9 megabytes. Na Figura 12 é possível ver um exemplo de imagem extraída do microscópio.

Figura 12 – Imagem de clínquer rico em  $C_3S$  (cristais azulados).



Fonte: Autor (2022).

#### **4.4.1 Pré-processamento geral**

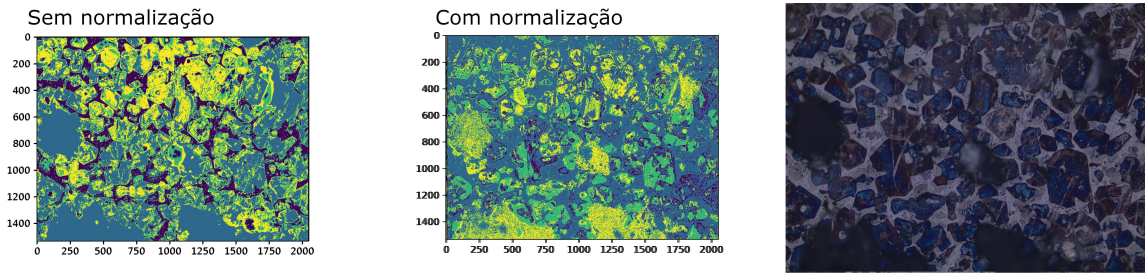
Esta etapa consiste, basicamente, na aplicação de filtros para amenizar ruídos e evidenciar características de interesse na imagem. Neste trabalho, para obtenção de resultados preliminares, é usado o filtro bilateral (PARIS *et al.*, 2009) para suavização e redução de ruído, preservando as bordas dos componentes mineralógicos representados na mesma. Seu funcionamento se dá pela substituição de cada pixel por um peso médio de seus pixels vizinhos. Após essa etapa, é aplicado o mapeamento de cor para facilitar a distinção entre os componentes na imagem, utilizando como principal referencial o nível de saturação da imagem. Finalmente, a imagem passa pela normalização para facilitar a etapa de segmentação.

#### **4.4.2 Segmentação com K-Means**

Em geral, os algoritmos de segmentação agrupam os dados levando em consideração, além de outras similaridade, a distância dos pixels vizinhos. Assim, inicialmente, os dados são normalizados. Pois, caso a base não esteja normalizada, ou seja padronizada, o algoritmo tende a calcular erroneamente, uma vez que os dados podem estar em escalas muito diferentes.

O K-Means é o algoritmo utilizado para agrupamento nesta exploração. Este foi inicialmente proposto por **Stuart Lloyd**. Este algoritmo agrupa os pixels com base na sua similaridade com pontos inicialmente propostos, os quais são denominados centroides (ou *cluster* no inglês). Estes pontos, são definidos com base no conhecimento prévio que o analista tem do

Figura 13 – Comparação da aplicação do K-Means com a imagem normalizada e não normalizada. A mais à direita é a original sem processamento.



Fonte: Autor (2022).

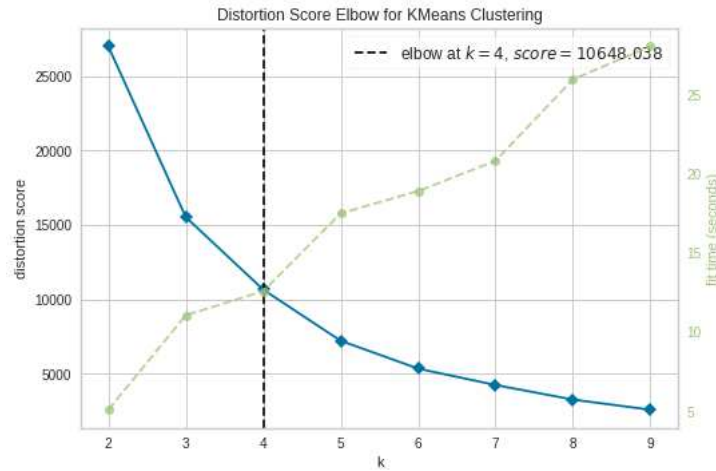
problema. Assim, caso o objetivo seja separar quatro elementos da imagens, devem ser propostos pelo menos quatro centroides. Esse é um hiperparâmetro do algoritmo que é estimado aqui com base na verificação de desempenho. Dessa forma, o melhor número de centroide será aquele que melhor elevar a eficácia da rede neural (CELEBI, 2014).

O principal intuito nesta etapa é usar a estratégia de divisão e conquista, ou seja, dividir o problema em partes menores com o objetivo de facilitar a busca dos componentes de forma individual. Assim, o algoritmo K-Means segmenta a imagem gerando quatro imagens (grupos ou classes) com cada uma possuindo componentes similares. Na Figura 14 é possível visualizar um gráfico que mostra o melhor parâmetro para a quantidade de clusters do K-Means. Este método, denominado em inglês como *KElbow*, testa o parâmetro, dado um intervalo, de forma iterativa. Em cada iteração, é verificada a soma obtida dos quadrados das distâncias de cada ponto em relação ao seu centróide. Com isso, é calculado o escore de distorção pela fórmula  $\sum_1^n distance(P_i, C_1)$ , onde  $P_i$  é o ponto analisado,  $C_1$  o centróide e *distance* a distância entre os dois pontos considerados. O escore de distorção indica o nível de similaridade entre os dados do mesmo centroide. No gráfico é possível observar um ponto onde ambas as linhas se encontram, este ponto é considerado o melhor parâmetro para o K-Means.

#### 4.4.3 Segmentação com *K-Nearest-Neighbors*( *KNN*)

O algoritmo *K-Nearest-Neighbors* (K Vizinhos mais próximos) (KRAMER, 2013) tem similaridade com o K-Means. Porém, o primeiro contém o diferencial de ser supervisionado. Assim, esse algoritmo pertence ao grupo de algoritmos de classificação. O seu funcionamento se dá por meio do cálculo das distâncias (sendo o cálculo da distância euclidiana o mais comumente utilizado) dos dados em relação aos seus vizinhos, e em seguida os classifica. Com isso, é possível mapear os componentes químicos presentes na imagem de forma que a classe nunca

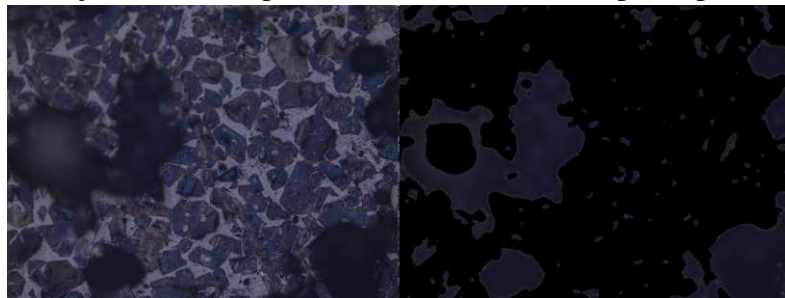
Figura 14 – Gráfico de distorção mostrando a performance do K-Means de acordo com seu parâmetro de quantidade de centróides.



Fonte: Autor (2022).

mude de posição, ou seja, se o cristal de  $C_3S$  foi treinado para ser classificado como da classe 0, este sempre será encontrado na camada 0 da imagem gerada pelo algoritmo. Note que com o K-Means não é possível fazer isto, pois trata-se de um algoritmo não supervisionado, fazendo com que as classes mudem de rótulo de uma imagem para outra. Contudo, o treinamento do KNN se dá pela utilização da imagem do clínquer pré-processada e da saída do K-Means, uma vez que ele já irá gerar uma imagem segmentada com as classes definidas, restando assim que o KNN apenas aprenda a "imitar" o K-Means. Com isso, o KNN calcula pixel a pixel bastando apenas uma imagem para efetuar seu treinamento e o mesmo ser capaz de segmentar outras imagens de forma efetiva. Na Figura 15 é possível observar exemplos das segmentações executadas por ambos os algoritmos. A imagem a esquerda da figura mostra a segmentação apenas pelo K-means e a imagem da direita a mesma imagem segmentada pelo KNN.

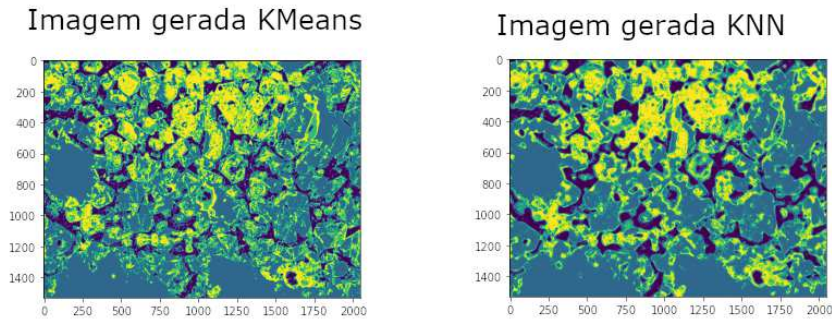
Figura 15 – Segmentação executada pelo KNN. À direita são os poros presentes na imagem.



Fonte: Autor (2022).



Figura 16 – Segmentações realizadas pelo K-Means e KNN.



Fonte: Autor (2022).

#### 4.4.4 Identificação dos contornos

Na identificação dos contornos, cada imagem gerada pelo K-Means é binarizada de acordo com um intervalo numérico (50,255), ou seja, qualquer pixel abaixo do intervalo será transformado em 0 e valores dentro do intervalo 255. A imagem binarizada é então repassada para a função responsável por encontrar os contornos de cada componente na imagem, "*findContours*". Com isso, cada componente é recortado de cada imagem gerada pelo KMeans resultando em uma imagem de cada componente isoladamente.

### 4.5 Classificação com CNN

Como citado anteriormente, a rede neural convolucional pode ser utilizada para classificação da imagem, com a tarefa de identificar os componentes que deverão ser analisados no passo posterior. Aqui, os elementos tratados são poros e  $C_3S$ , mas outros elementos podem ser encontrados como CaO livre, magnésio,  $C_2S$  etc. Os elementos que aparecem na imagem e não são foco das análises são descartados, que em geral são componentes pequenos ou fazem parte de outros maiores que são gerados na fase de identificação dos contornos. As seguintes seções descreveram sobre o pré-processamento, arquitetura e treinamento da CNN.

#### 4.5.1 Pré-processamento da base de dados de treinamento para a CNN

Os dados de entrada da rede são imagens na escala HSV, os quais possuem três canais (matiz, saturação e valor de luminosidade) de tamanho 50x50 (Largura x Altura). Assim, a entrada final tem dimensões 50x50x3. Essas dimensões foram escolhidas devido as imagens

da base de dados possuírem tamanhos diferentes entre si, fazendo com que fosse necessário a escolha de um tamanho de forma que a perda de informação seja mínima. Para efetuar o redimensionamento das imagens é utilizada a técnica de interpolação bicúbica, técnica que pode ser vista em Han (2013), a qual retorna bons resultados. As imagens na base de dados podem chegar ao tamanho de 10x10 e outras até 500x200. Aumentar uma imagem pequena para se adequar ao tamanho médio das imagens acima de 100x100 resultaria em uma imagem com qualidade inferior, pois o algoritmo cria pixels artificialmente utilizando os seus vizinhos como referência, com isso, quanto maior a imagem mais é perceptível "borrões". Um exemplo disso pode ser observado na Figura 17. Por fim, é aplicado o filtro mediano (em inglês *median filter*)(CHEN *et al.*, 1999) para remoção de ruído e uniformização das cores na imagem. Esse procedimento é realizado em conjunto com a conversão para o padrão de cores HSV. No treinamento exploratório foram utilizadas 1504 imagens, dessas, 70% foi utilizado para treino e o restante para testes, afim de se obter uma média geral de sua precisão.

Figura 17 – Exemplo do aumento de uma imagem de Magnésio, no centro(destacado em vermelho) o seu tamanho original 10x13 e ao fundo o tamanho aumentado de 100x100.



Fonte: Autor (2022).

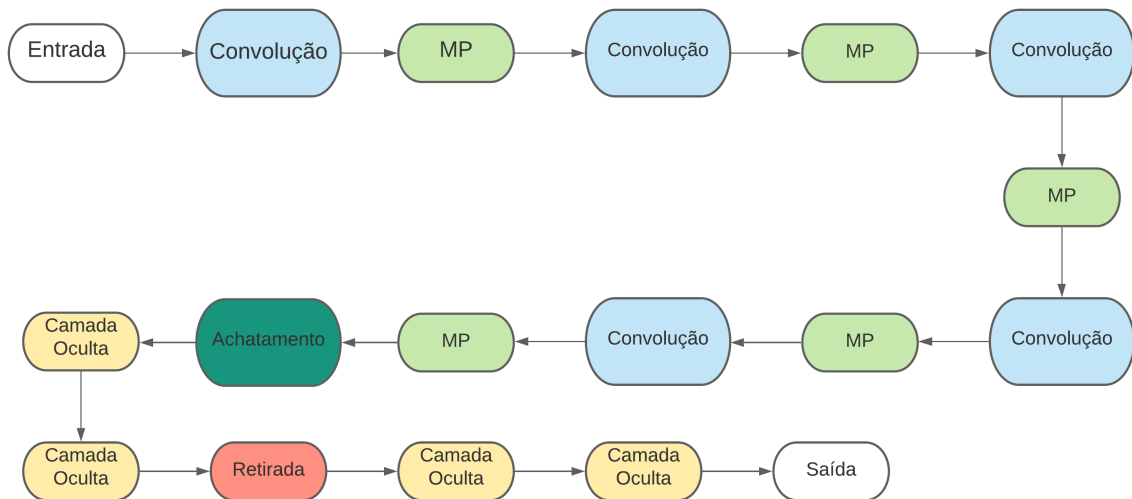
#### 4.5.2 Arquitetura da CNN

A arquitetura da CNN pode ser dividida em duas etapas, a convolucional e a classificação. A etapa convolucional começa na camada de entrada, seguida por cinco camadas convolucionais. Entre cada camada há uma camada denominada *max pooling* (referenciado como MP na Figura 18), que recebe a saída da convolução e filtra novamente a imagem seguindo a regra de considerar apenas os maiores valores, descartando os menores, e assim retorna apenas

as informações mais importantes.

A etapa de classificação se inicia na camada denominada de "achatamento", a qual converte cada matriz para um vetor unidimensional e passa para as camadas ocultas. No total são quatro, cuja ativação é a ReLU para todos os neurônios. No centro das camadas ocultas há a camada de retirada (mais conhecida em inglês como *dropout*), que tem como função descartar aleatoriamente porcentagem dos dados da rede, impedindo que esses participem do processo de treinamento. Essa camada tem a finalidade de dificultar que a rede neural fique superajustada (do inglês *overfitting*). Isto faz com que a rede retorne bons resultados no treino, mas falhe na etapa de teste. Além disto, essa camada também aumenta a precisão da rede utilizando uma base de dados menor. Por fim, a camada de classificação aplica a função de *softmax* (também conhecida como *softargmax* ou função exponencial normalizada) para identificar a qual componente pertence a imagem, retornando um vetor com a probabilidade entre 0 e 1 para cada posição. Assim, cada posição é um componente a ser identificado (poro,  $C_3S$ , descarte etc). A arquitetura pode ser melhor visualizada na Figura 18.

Figura 18 – Arquitetura visual da rede neural convolucional.



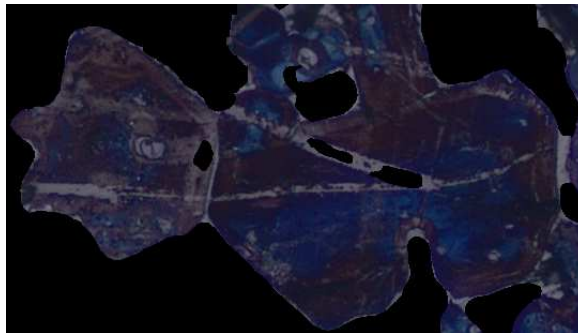
Fonte: Autor (2022).

Com isso, a exploração mostrou que a segmentação é possível até mesmo com os algoritmos citados acima. Porém, existe um problema, devido a aplicação do filtro da mediana para remoção de ruído e uniformização de cores, o que acaba aproximando componentes próximos, os transformando em parte do  $C_3S$ . O resultado da segmentação em relação ao  $C_3S$  é uma imagem com os cristais colados, o que dificulta a aplicação de análises posteriores como checagem de forma e contagem de cristais. Isso pode ser visualizado na Figura 19. Além disso,

o algoritmo apresentado na exploração erra em média 10% dos pixels na imagem, sendo estes pixels outros componentes ou a ausência dos pixels de um cristal de  $C_3S$ . Isto acontece devido a dificuldade na identificação do  $C_3S$  que não apresenta uma forma exata definida e sua coloração que varia em tom de azul e marrom quando está em transição para  $C_2S$ . Um possível algoritmo que pode apresentar melhor performance para esse problema é a MASK R-CNN (HE *et al.*, 2017), um tipo de rede neural convolucional mais moderna que classifica e localiza os objetos de interesse na imagem, juntamente com esse algoritmo pode ser utilizado o filtro bilateral (PARIS *et al.*, 2009) de forma a remover ruído e uniformizar as cores preservando as bordas dos objetos na imagem.

Toda a exploração das imagens foi feita em linguagem *Python* em sua versão 3 (ROSSUM; DRAKE, 2009). As principais bibliotecas utilizadas são: *opencv* (BRADSKI, 2000), para técnicas de visão computacional; *scikit-learn* (PEDREGOSA *et al.*, 2011), para utilização do K-Means, KNN, *KElbow* e normalização, e *tensorflow* (ABADI *et al.*, 2016) para criação de redes neurais.

Figura 19 – Cristais de  $C_3S$  colados por componente esbranquiçado.



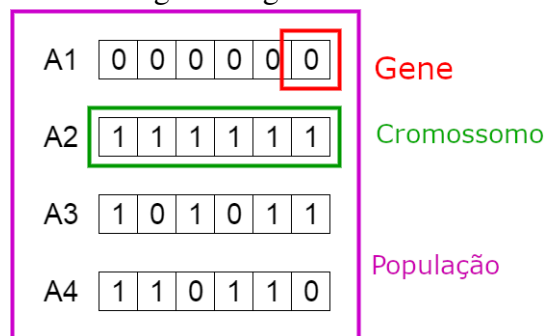
Fonte: Autor (2022).

#### 4.6 Algoritmo Genético

O algoritmo genético (MIRJALILI, 2019) é uma heurística de busca inspirada na teoria da evolução natural de Charles Darwin. Esse algoritmo reflete o processo de seleção natural onde os indivíduos mais aptos são selecionados para reprodução a fim de produzir descendentes da próxima geração. O processo de seleção natural começa com a seleção dos indivíduos mais aptos de uma população. Eles produzem descendentes que herdam as características dos pais e serão adicionados à próxima geração. Se os pais tiverem melhor aptidão, seus filhos serão melhores que os pais e terão mais chances de sobreviver. Esse processo continua iterando e, no final, uma geração com os indivíduos mais aptos será encontrada. É possível utilizar esse

processo como uma heurística para um algoritmo de busca. Esse algoritmo pode ser dividido em cinco etapas de funcionamento. A primeira etapa é a geração da população inicial onde cada indivíduo dessa população é uma solução para o problema a ser resolvido pelo algoritmo. Esses indivíduos possuem os chamados genes que são um conjunto de variáveis que em conjunto são conhecidas como cromossomo (geralmente representado como um conjunto de zeros e uns). É possível visualizar uma estrutura mais visual desses dados na Figura 20.

Figura 20 – Estrutura dos dados no algoritmo genético.



Fonte: Autor (2022).

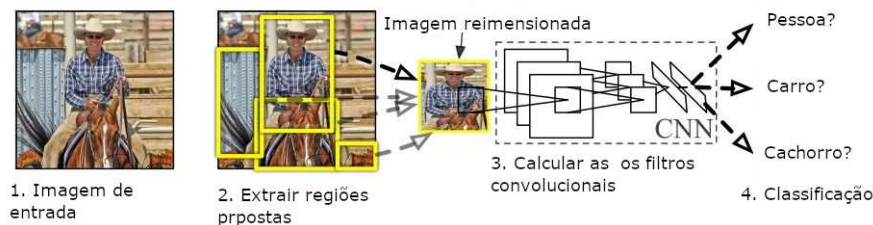
A segunda etapa é a função de aptidão (mais conhecida no inglês como *Fitness Function*) que determina o escore de aptidão que indica o quão boa é uma solução (indivíduo) para a solução do problema. Essa função é um dos parâmetros passados para o algoritmo. A terceira etapa é a de seleção, onde dois pares de indivíduos com os melhores escores de aptidão são selecionados para a próxima iteração do algoritmo. A quarta etapa é a de cruzamento, cada par de indivíduos são combinados a partir de um ponto dentro de seus genes gerando um indivíduo filho. Existem vários métodos para o cruzamento, um exemplo é escolher aleatoriamente um ponto de cruzamento de dentro dos genes. A última etapa é a de mutação, com os indivíduos filhos gerados seus genes podem ser selecionados aleatoriamente para serem modificados (um 0 pode se tornar 1 por exemplo). A mutação ocorre para manter a diversidade dentro da população e evitar a convergência prematura. O algoritmo termina se a população convergiu (não produz descendentes significativamente diferentes da geração anterior).

#### 4.7 Faster R-CNN

A Faster R-CNN (do inglês *Faster Region Convolutional Neural Network*) (REN *et al.*, 2015) é um tipo de rede neural convolucional de última geração mais utilizada da família R-CNN. Esta rede neural é uma melhoria da Fast R-CNN (do inglês *Fast Region Convolutional*

*Neural Network*) ambas usadas no problema de detecção de objetos utilizando caixa delimitadora. Essa caixa delimitadora é a responsável por localizar o objeto de interesse (objeto da qual a rede está sendo treinada para identificar) e delimitar sua área em um formato de retângulo, ou seja, acrescenta um retângulo que encapsula o objeto na imagem. O diferencial da Faster R-CNN é prever com precisão e rapidez as localizações de diferentes objetos. Isso ocorre por meio da adição de uma ramificação responsável por gerar propostas de regiões (conhecida como RPN do inglês *Region Proposal Network*) na imagem que são alimentadas ao modelo de detecção (Fast R-CNN) para inspecionar objetos, ou seja, essa ramificação recorta áreas da mesma imagem que possuem a maior chance de possuir os objetos de interesse para a classificação. O funcionamento da Faster R-CNN é composto por três etapas: a RPN gera propostas de regiões usando o algoritmo de Busca Seletiva, cada imagem é redimensionada para um tamanho fixo pré-definido e então é extraído um vetor de características de cada região de proposta e por fim um algoritmo SVM pré-treinado classifica a região de proposta entre plano de fundo e uma das classes de objetos de interesse. É possível visualizar o funcionamento da FASTER R-CNN mais simplificada na Figura 21.

Figura 21 – Funcionamento da Faster R-CNN



Fonte: autor

#### 4.8 Mask R-CNN

A Mask R-CNN (do inglês *Mask Region Convolutional Neural Network*) (HE *et al.*, 2017) é um tipo de rede convolucional que usa a Faster R-CNN (REN *et al.*, 2015) adicionando uma ramificação para prever máscaras de segmentação em cada Região de Interesse (ROI do inglês *Region of Interest*) em paralelo a ramificação para classificação e regressão de caixa delimitadora. É possível visualizar um exemplo dessa caixa e da máscara na Figura 22 onde a caixa é o retângulo em azul e a máscara a área pintada em azul no interior da caixa.

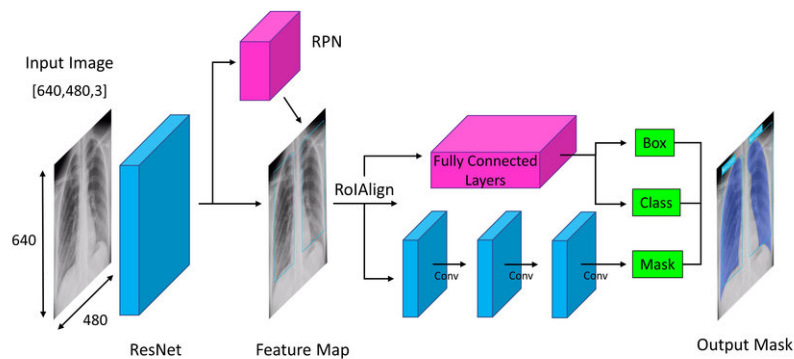
Figura 22 – Exemplo de máscara e caixa delimitadora



Fonte: autor

A camada responsável por gerar a máscara é uma pequena camada de neurônios completamente conectada (FCN do inglês *Fully Connected Network*) aplicada para cada ROI, prevendo uma máscara de segmentação correspondente a sua classe para cada pixel que faz parte da ROI. Com essa rede neural é possível aplicar a segmentação semântica que é definida como uma divisão com base na identificação de cada objeto de interesse presente na imagem. Na Figura 23, é possível visualizar a arquitetura da Mask R-CNN para a segmentação de pulmões.

Figura 23 – Arquitetura da MASK R-CNN.

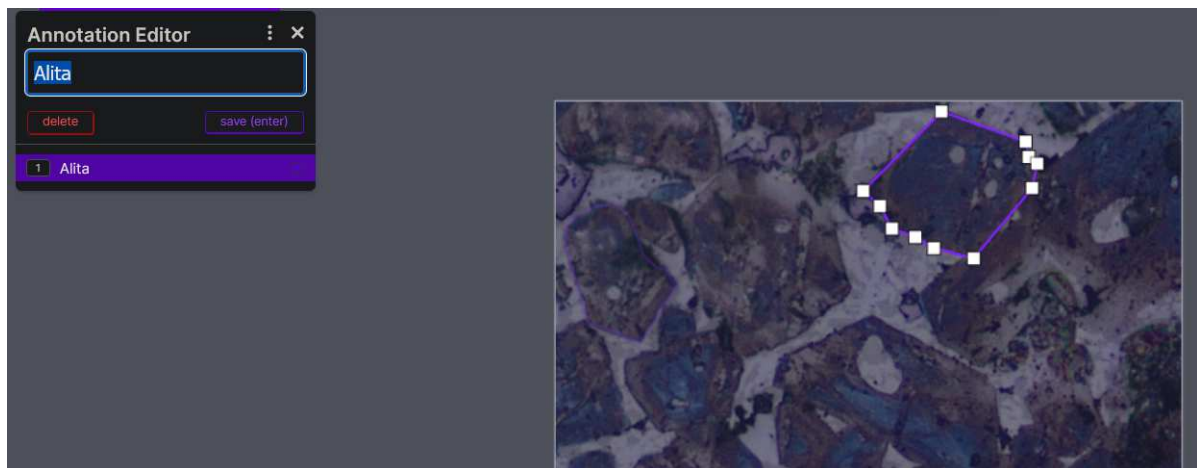


Fonte: (PODDER *et al.*, 2021)

Com este método mais robusto é possível aplicar a segmentação semântica nas imagens de clíquer afim de localizar, gerar a máscara e identificar o  $C_3S$ . Para efetuar o treinamento da Mask R-CNN é necessário construir a base de dados de treino e teste. Para essa rede, são necessárias as imagens dos objetos de interesse e um arquivo de texto que contenha as informações de cada máscara presentes em cada imagem da base de dados (localização, coordenada dos pontos que formam a máscara e etc). As principais informações desse arquivo

são as coordenadas dos pontos que formam as máscaras que englobam os objetos de interesse e a suas respectivas classes. Por meio de um sistema de anotação de máscaras e caixas delimitadoras em imagens é possível criar esse arquivo facilmente, na Figura 24 é possível visualizar um exemplo desse sistema (nesta pesquisa foi utilizado o Roboflow (2022)). Com esta base de dados concluída é possível treinar a Mask R-CNN por meio de transferência de aprendizagem (TORREY; SHAVLIK, 2010) com os pesos pré treinados da COCO (do inglês *Common Objects in Context*) que trata-se de pesos de uma Mask R-CNN já treinados anteriormente para segmentação e classificação de diversos objetos, animais e etc. A transferência de aprendizagem apresenta as vantagens de melhorar a acurácia dos resultados e reduzir o tempo de treinamento.

Figura 24 – Exemplo de um sistema para anotação de máscara, é possível ver um polígono (máscara) que engloba o cristal de  $C_3S$  a direita e a esquerda na caixa de texto a sua classe.



Fonte: Autor (2022).



## 5 RESULTADOS

Neste capítulo será abordado o treinamento da Mask R-CNN e seus resultados gerados, assim como técnicas para melhoria de performance e métricas para indicar a eficiência do modelo. Por fim, os resultados finais desse estudo e sugestões de melhorias.

### 5.1 Treinamento da Mask R-CNN

Para treinamento da Mask R-CNN foi utilizado uma base de dados com 85 imagens, sem pré processamento, sendo dessas 56 imagens 29 para treino, 12 para teste e 15 para validação. Além disso, para esse treinamento inicial, foram selecionados apenas cristais com formas mais regulares possível e inclusão um novo componente que é o poro. Essa componente nada mais é que outras componentes que foram queimadas durante a fabricação do clínquer e o resultado disso é uma mancha preta. Essa componente foi incluída para evitar que a rede neural classifique alguns cristais de  $C_3S$  erroneamente como poro, pois alguns poros possuem um tom de azul escuro.

Figura 25 – Comparação entre a imagem segmentada manualmente (correta) e a calculada pela Mask R-CNN.

Segmentação correta

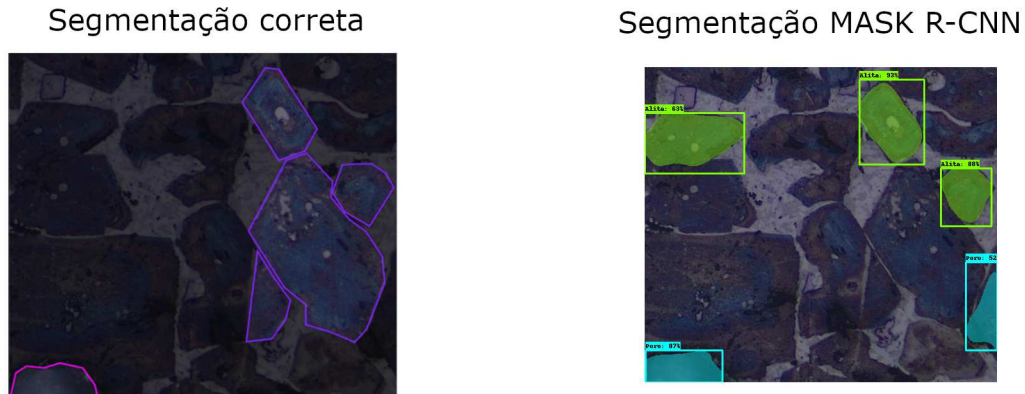


Segmentação MASK R-CNN



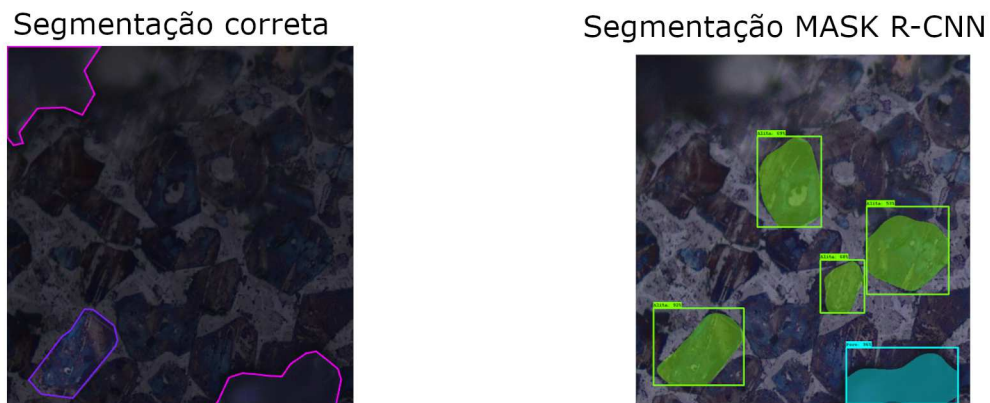
Fonte: Autor (2022).

Figura 26 – Cristal de  $C_3S$  classificado erroneamente como poro (componente em azul mais a direita) na imagem segmentada pela MASK R-CNN.



Fonte: Autor (2022).

Figura 27 – Poro mais acima não foi identificado pela MASK R-CNN.



Fonte: Autor (2022).

Na Figura 25 é possível notar que a rede identificou um cristal com boa forma que não foi incluído na segmentação manual (feita por um humano) porém classificou erroneamente uma componente como poro (componente em azul mais a esquerda). Além de ter aumentado a máscara do poro mais a direita. Cenários parecidos podem ser observados nas Figuras 26 e 27.

## 5.2 Treinamento da Mask R-CNN com Filtro Bilateral

O treinamento foi repetido mas dessa vez utilizando o filtro bilateral (PARIS *et al.*, 2009) com os parâmetros 2, 52 e 84 que foram descobertos utilizando algoritmo genético (MITCHELL, 1995). Esse algoritmo mimetiza o processo biológico de evolução afim de solucionar problemas e possui uma série de parâmetros, mas os mais importantes são a estrutura de dados que o algoritmo genético deve calcular. Para descobrir os melhores parâmetros que

melhorem a acurácia da Mask R-CNN via filtro bilateral é utilizado um conjunto de 3 números reais que representam os 3 parâmetros do filtro bilateral: o diâmetro de cada vizinhança de pixel que é usada durante a filtragem, sigma no espaço de cores onde um valor maior do parâmetro significa que cores mais distantes na vizinhança do pixel são misturadas e sigma no espaço de coordenadas onde um valor maior do parâmetro significa que os pixels mais distantes irão influenciar uns aos outros desde que suas cores estejam próximas o suficiente. O segundo parâmetro importante para o algoritmo genético é a função que avalia o treinamento, retornando valores que ajudem o algoritmo genético a progredir. Essa função retorna valores entre o infinito negativo e 0, onde quanto mais abaixo de 0 indica que o algoritmo está longe da solução.

---

**Algoritmo 1:** Função utilizada para gerar o escore de aproximação da solução para o algoritmo genético.

---

```
function calculate error (masksvalidated, maskspredicted);
Input : 2 matrizes com valores binários (0 ou 1), onde masksvalidated e
          maskspredicted correspondem as máscaras para validação e máscaras geradas
          pela Mask R-CNN respectivamente.
Output : Um número real entre 0 e infinito negativo.
areaerrors = [];
for index = 1, 2, ... length(masksvalidated) do
    validarea = sum(masksvalidated[index]);
    calculatedarea = sum(maskspredicted[index]);
    areaerrors.add(calculatedarea / validarea);
end
return (1 - mean(areaerrors)) * -1;
```

---

No Algoritmo 1 é possível visualizar o pseudo código dessa função. Essa recebe dois conjuntos de máscaras onde uma delas são as máscaras geradas manualmente (por um humano) e as máscaras geradas pela Mask R-CNN, com ambas possuindo a mesma quantidade de elementos. Para cada máscara é calculado a sua área é feito uma divisão da área da máscara calculada pela Mask R-CNN e a área da máscara de validação. O resultado dessa divisão indica a similaridade dessas máscaras onde quanto mais próximo de 1 indica alta acurácia da Mask R-CNN para a inferência. Ao fim da função é retornado a média desses valores calculados subtraído de 1 e multiplicado por -1. Esse cálculo serve para penalizar o algoritmo genético caso os valores sejam muito distantes de 1, e assim provocar que o algoritmo busque valores mais satisfatórios.

O treinamento foi executado novamente, mas dessa vez utilizando o filtro bilateral em cada imagem da base de dados de treino ao ser alimentada para a rede (as mesmas imagens foram

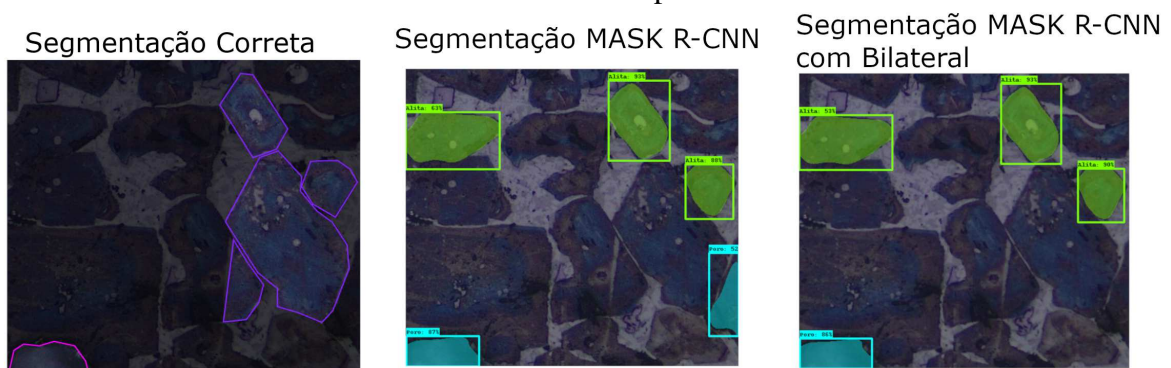
utilizadas nesse treinamento). Nas Figuras 28, 29, 30 é possível visualizar a diferença causada pelo filtro bilateral nas inferências da Mask R-CNN. O filtro bilateral melhorou a identificação dos cristais de  $C_3S$  na base de dados de treino, além disso evidenciou a diferença entre os cristais e os poros. Nas Figuras 25 e 26 foi evidenciado a classificação errônea de componentes como poro, o que foi solucionado com o filtro bilateral. Isso pode acontecer devido ao poro nem sempre estar numa escala pura de preto, podendo variar para um azul escuro se assimilando a principal característica do  $C_3S$ .

Figura 28 – Comparação de resultados entre as imagens de segmentação, nesta figura é possível notar que a Mask R-CNN identificou um cristal diferente que estava mais regular com o filtro bilateral.



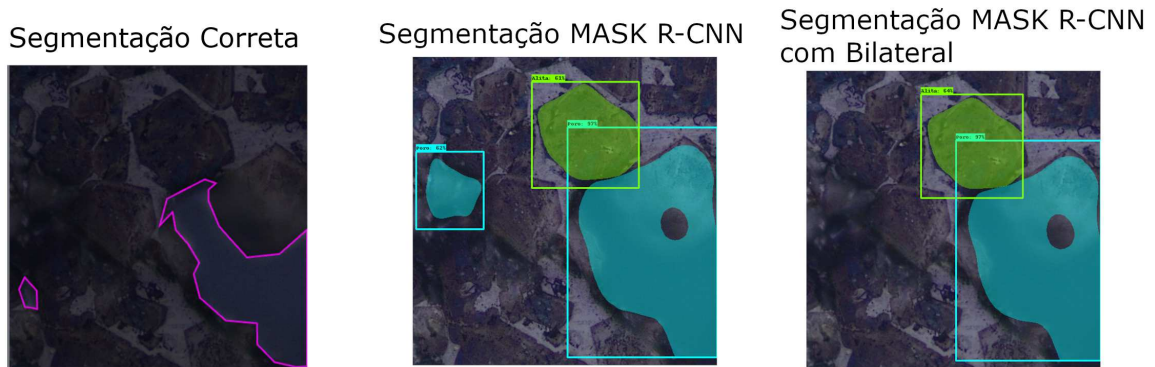
Fonte: Autor (2022).

Figura 29 – Nessa figura é possível notar que na imagem com filtro bilateral um cristal de  $C_3S$  não foi classificado erroneamente como poro.



Fonte: Autor (2022).

Figura 30 – Nessa figura é possível notar que a componente mais a esquerda da imagem não foi classificada erroneamente como poro.



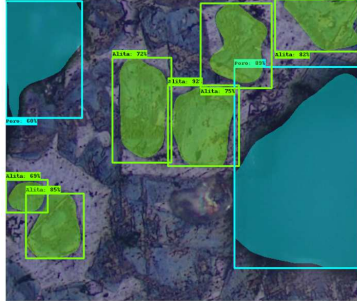
Fonte: Autor (2022).

### 5.3 Treinamento final da Mask R-CNN

Nessa etapa foi realizado o treinamento utilizando uma base de dados com 300 imagens com zoom e iluminação diferentes para o treinamento da Mask R-CNN novamente e com a utilização do filtro bilateral. Os resultados desse treinamento se mostraram mais satisfatórios, porém, notou-se a presença de erros na seleção de cristais de  $C_3S$ , principalmente na não inclusão de alguns cristais. É possível visualizar alguns dos resultados nas Figuras 31 até 36.

Figura 31 – Nessa figura é possível notar que o cristal de  $C_3S$  mais abaixo e ao lado do poro não foi identificado.

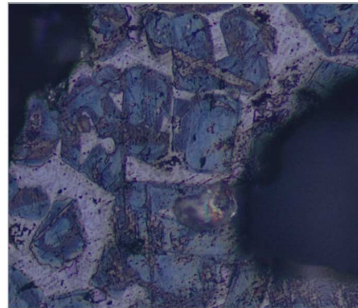
Imagem segmentada pela  
MASK R-CNN



Segmentação correta



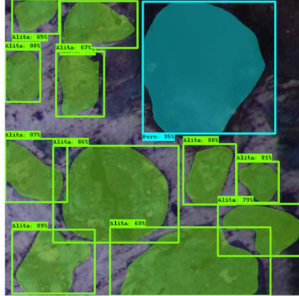
Imagem Original



Fonte: Autor (2022).

Figura 32 – Nessa figura é possível notar bons resultados e apenas um erro na não identificação de 2 cristais de  $C_3S$  ao lado esquerdo do poro.

Imagem segmentada pela  
MASK R-CNN



Segmentação correta

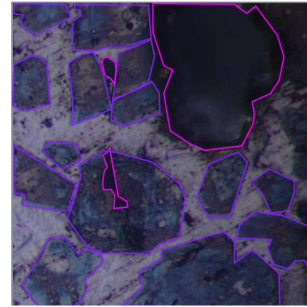
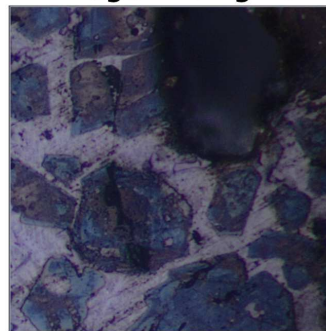


Imagem Original



Fonte: Autor (2022).

Figura 33 – Nessa figura é possível notar que o maior cristal de  $C_3S$  mais a direita foi identificado porém o cristal a sua esquerda não foi, além disso, a máscara do poro está menor do que deveria.

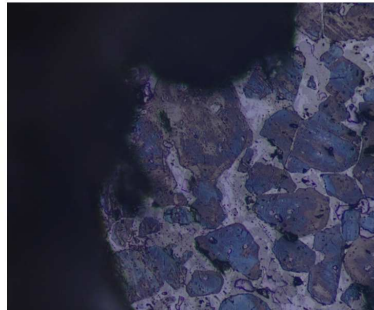
Imagem segmentada pela MASK R-CNN



Segmentação correta



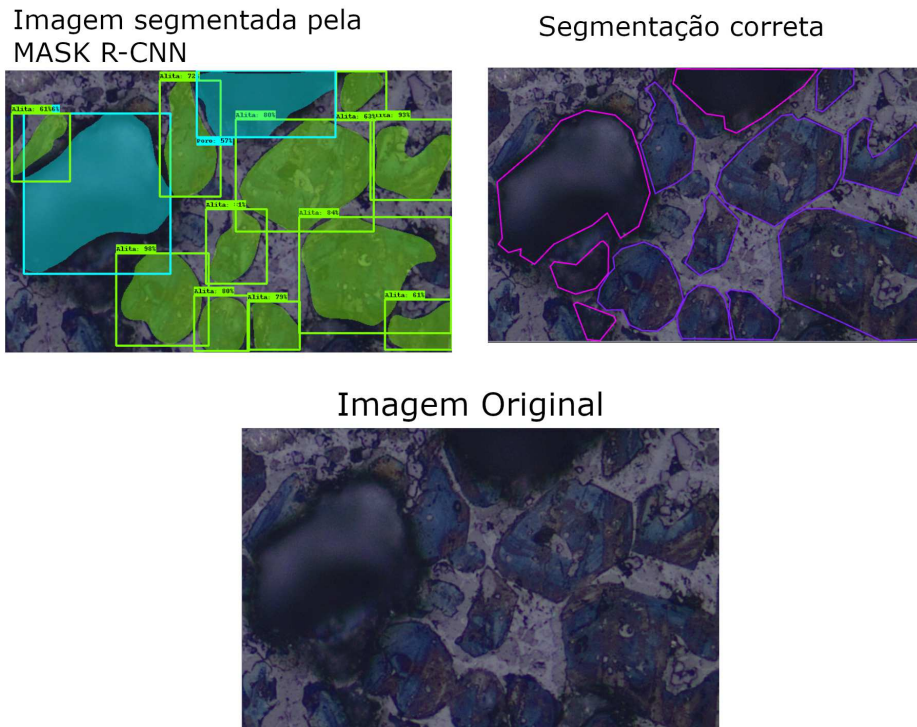
Imagem Original



Fonte: Autor (2022).

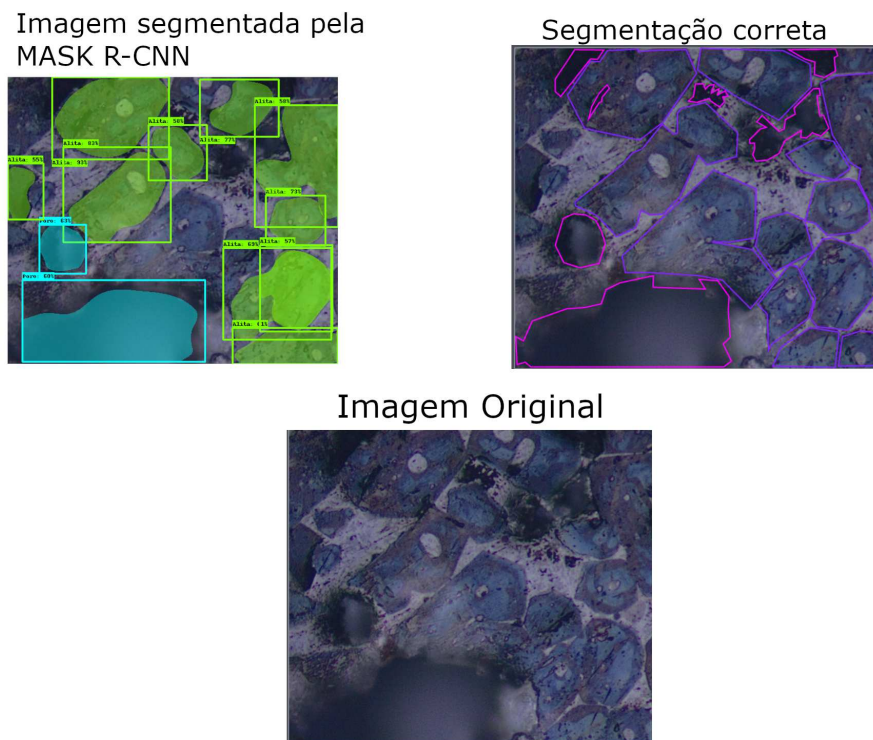


Figura 34 – Nessa figura é possível notar ótimos resultados.



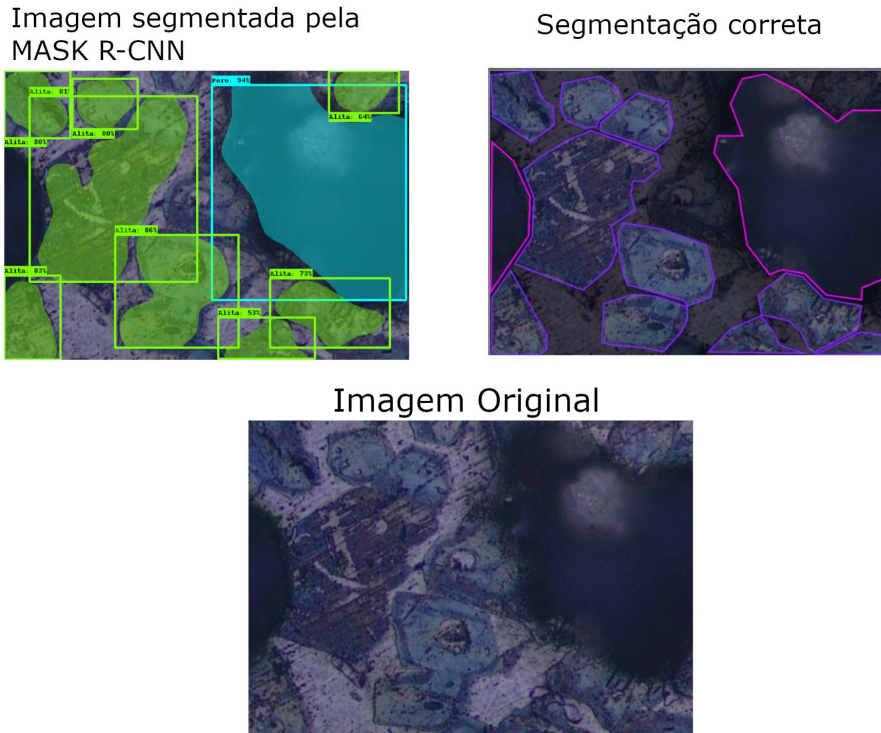
Fonte: Autor (2022).

Figura 35 – Nessa figura é possível notar o cristal de  $C_3S$  acima do poro mais a baixo da imagem não identificado.



Fonte: Autor (2022).

Figura 36 – Nessa figura houveram erros apenas na separação de cristais de  $C_3S$  muito grudados o que pode ser analisado juntamente com um especialista na área como esse cenário seria tratado.



Fonte: Autor (2022).

Na Tabela 2 é possível visualizar a validação das classificações geradas pela Mask R-CNN. Há 4 colunas onde duas se referem as identificações calculadas pela Mask R-CNN e as outras duas de validação. É possível notar que a Mask R-CNN não conseguiu identificar todas as componentes presentes nas imagens. Entretanto, grande parte das imagens utilizadas para o seu treinamento eram de qualidade inferior providas de cimento antigo e foram utilizadas apenas devido a baixa quantidade de imagens de qualidade. Por esse fato, pode ter ocasionado sua baixa precisão. Somado a isso, ocorreu pouca interação dos especialistas em microscopia do clínquer durante o presente estudo o que dificultou o completo entendimento dos cenários possíveis (como tratar diferentes tons de cor dos poros, quais os critérios de escolha para cristais muito unidos entre outros cenários) para o treinamento adequado da Mask R-CNN. Contudo, é notório a identificação de bons cristais de  $C_3S$  assim como poros, além disso, não ocorreu erros de classificações (poro ser classificado como  $C_3S$  e vice versa).

Tabela 2 – Comparação da monografia com os trabalhos relacionados

<b>Classes</b>	<b>Precisão</b>	<b>Recall</b>	<b>F1-Score</b>	<b>Suporte</b>
<b>Erro</b>	0.00	0.00	0.00	3
<b>Alita</b>	0.98	0.57	0.72	303
<b>Poro</b>	1.00	0.40	0.57	157

Fonte: elaborado pelo autor (2022).

## 6 CONCLUSÕES E TRABALHOS FUTUROS

O desenvolvimento do presente estudo possibilitou uma análise de como o processo de microscopia do clínquer é complexa e custoso, possuindo várias problemáticas. Esse processo possui um manual com várias normas que informam pontos positivos ou negativos durante todo o processo de fabricação do cimento. O especialista em análise de clínquer em sua análise utiliza esse manual, seu conhecimento em química e seu tempo de experiência na busca de características de cada componente presente na imagem, afim de produzir o relatório de qualidade da análise do clínquer.

Com o que foi apresentado no Capítulo 5 a Mask R-CNN, que é um algoritmo de redes neurais convolucionais, foi treinada com uma base de dados de imagens de diversas fontes diferentes da mesma planta cimenteira. Entretanto, grande parte dessas imagens foram obtidas de clínquer antigo do qual não seria usado para uma microscopia do clínquer por um analista. Aliado a isso, ocorreu pouca interação de especialistas em microscopia do clínquer afim de validar os resultados da Mask R-CNN e sanar dúvidas em relação as formas dos cristais de  $C_3S$  que deviam ser identificados. Mesmo com a presença desses problemas, a Mask R-CNN aliado ao filtro bilateral foi capaz de identificar uma quantidade satisfatória de cristais  $C_3S$  nas imagens, e aliado ao filtro bilateral impediu a classificação incorreta de cristais de  $C_3S$  e poros. Com isto, é possível afirmar que a identificação e segmentação de  $C_3S$  por meio de visão computacional é uma solução possível de ser implementada.

Porém, afim de se atingir resultados comparáveis ao de um analista de clínquer, devem ser utilizadas imagens de qualidade e de preferência a presença de um especialista em microscopia do clínquer para melhor entendimento do processo, assim como a identificação correta das características das componentes de interesse presentes nas imagens.

## REFERÊNCIAS

- ABADI, M.; BARHAM, P.; CHEN, J.; CHEN, Z.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; IRVING, G.; ISARD, M. *et al.* Tensorflow: A system for large-scale machine learning. In: **12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)**. [S.l.: s.n.], 2016. p. 265–283.
- BOHÁČ, M.; KHONGOVÁ, I.; ZEZULOVÁ, A.; KREJČÍ, M. K.; KUBÁTOVÁ, D.; STANĚK, T. The role of cuo on the microstructure and phase composition of so3-activated clinker. **Journal of Microscopy**, Wiley Online Library, 2021.
- BRADSKI, G. The OpenCV Library. **Dr. Dobb's Journal of Software Tools**, 2000.
- BRE, F. **MULTILAYERIMG**. 2017. Acessado: 10-01-2022. Disponível em: <[https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o\\_fig1\\_321259051](https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051)>.
- CANNY, J. A computational approach to edge detection. **IEEE Transactions on pattern analysis and machine intelligence**, Ieee, n. 6, p. 679–698, 1986.
- CELEBI, M. E. **Partitional clustering algorithms**. [S.l.]: Springer, 2014.
- CHEN, T.; MA, K.-K.; CHEN, L.-H. Tri-state median filter for image denoising. **IEEE Transactions on Image processing**, IEEE, v. 8, n. 12, p. 1834–1838, 1999.
- COMMONS, C. **ConvImg**. 2019. [https://www.researchgate.net/figure/Depiction-of-the-convolution-layer-with-a-filter-in-convolutional-neural-network-CNN\\_fig3\\_38190342](https://www.researchgate.net/figure/Depiction-of-the-convolution-layer-with-a-filter-in-convolutional-neural-network-CNN_fig3_38190342). Acessado : 10 – 01 – 2022.
- CORTES, C.; VAPNIK, V. Support vector machine. **Machine learning**, v. 20, n. 3, p. 273–297, 1995.
- DHALLA, A. **perceptron**. 2021. Acessado: 10-01-2022. Disponível em: <<https://ai.plainenglish.io/the-rise-and-fall-of-the-perceptron-c04ae53ea465>>.
- GAO, S.; CAO, C.; HU, D.; XU, R.; CHENG, Y.; ZHU, M.; LI, S. Prediction of newhall navel orange internal quality based on digital microscopy. **International Journal of Agricultural and Biological Engineering**, v. 14, n. 6, p. 222–227, 2021.
- GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. [S.l.]: Chapman and Hall/CRC, 1995.
- GERMAND, A. Cimento portland pesquisa. 2017.
- HAN, D. Comparison of commonly used image interpolation methods. In: **Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)**. [S.l.: s.n.], 2013. v. 10.
- HE, K.; GKIOXARI, G.; DOLLAR, P.; GIRSHICK, R. Mask r-cnn. Oct 2017.
- HERNANDEZ, N.; LIZARAZO-MARRIAGA, J.; RIVAS, M. A. Petrographic characterization of portlandite crystal sizes in cement pastes affected by different hydration environments. **Construction and Building Materials**, Elsevier, v. 182, p. 541–549, 2018.

IGLESIAS, J. C. Álvarez; AUGUSTO, K. S.; da Fonseca Martins Gomes, O.; DOMINGUES, A. L. A.; VIEIRA, M. B.; CASAGRANDE, C.; PACIORNIK, S. Automatic characterization of iron ore by digital microscopy and image analysis. **Journal of Materials Research and Technology**, v. 7, n. 3, p. 376–380, 2018. ISSN 2238-7854. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2238785417308025>>.

JENA, M. C.; MISHRA, S. K.; MOHARANA, H. S. Application of industry 4.0 to enhance sustainable manufacturing. **Environmental Progress & Sustainable Energy**, v. 39, n. 1, p. 13360, 2019.

JOHNSON, D. **Backpropagation**. 2022. Acessado: 10-01-2022. Disponível em: <<https://www.guru99.com/backpropagation-neural-network.html>>.

JOURLIN, M.; PINOLI, J. C. Logarithmic image processing: the mathematical and physical framework for the representation and processing of transmitted images. **Advances in imaging and electron physics**, Elsevier, v. 115, p. 129–196, 2001.

JOURLIN, M.; ROUX, B.; FAURE, R.-M. Recognition of clinker phases by automatic image analysis. **Cement and Concrete Composites**, Elsevier, v. 23, n. 2-3, p. 207–214, 2001.

KALCHBRENNER, N.; GREFENSTETTE, E.; BLUNSOM, P. A convolutional neural network for modelling sentences. **arXiv preprint arXiv:1404.2188**, 2014.

KESHISHIAN, M.; AKBARI, H.; KHALIGHINEJAD, B.; HERRERO, J. L.; MEHTA, A. D.; MESGARANI, N. Estimating and interpreting nonlinear receptive field of sensory neural responses with deep neural network models. **Elife**, eLife Sciences Publications Limited, v. 9, p. e53445, 2020.

KIM, D. J.; LEE, C.-W.; SUH, Y.; JEONG, H.; JO, I.; MOON, J.; PARK, M.; WOO, Y. S.; HONG, B. H. Confocal laser scanning microscopy as a real-time quality-assessment tool for industrial graphene synthesis. **2D Materials**, IOP Publishing, v. 7, n. 4, p. 045014, 2020.

KRAMER, O. K-nearest neighbors. In: **Dimensionality reduction with unsupervised nearest neighbors**. [S.l.]: Springer, 2013. p. 13–23.

KRISHNAMURTHY, S.; BROWN, J. Q.; IFTIMIA, N.; LEVENSON, R. M.; RAJADHYAKSHA, M. Ex vivo microscopy: a promising next-generation digital microscopy tool for surgical pathology practice. **Archives of pathology & laboratory medicine**, the College of American Pathologists, v. 143, n. 9, p. 1058–1068, 2019.

LI, H.; ELMOATAZ, A.; FADILI, J. M.; RUAN, S. An improved image segmentation approach based on level set and mathematical morphology. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **Third international symposium on multispectral image processing and pattern recognition**. [S.l.], 2003. v. 5286, p. 851–854.

LUO, W.; LI, Y.; URTASUN, R.; ZEMEL, R. Understanding the effective receptive field in deep convolutional neural networks. In: **Proceedings of the 30th International Conference on Neural Information Processing Systems**. [S.l.: s.n.], 2016. p. 4905–4913.

MA, Z.; YU, H.; LU, S.; WANG, X.; LIU, H. Design and research of simulation system for pre-decomposition process of cement raw materials based on cloud platform. In: **2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC)**. [S.l.: s.n.], 2020. p. 594–598.

- MAROUŠEK, J.; MAROUŠKOVÁ, A.; KŮS, T. Shower cooler reduces pollutants release in production of competitive cement substitute at low cost. **Energy Sources, Part A: Recovery, Utilization, and Environmental Effects**, Taylor Francis, v. 0, n. 0, p. 1–10, 2020.
- MARTÍN-VIVEROS, J. I.; OLLÉ, A. Using 3d digital microscopy and sem-edx for in-situ residue analysis: A multi-analytical contextual approach on experimental stone tools. **Quaternary International**, Elsevier, v. 569, p. 228–262, 2020.
- MASOOD, T.; SONNTAG, P. Industry 4.0: Adoption challenges and benefits for smes. **Computers in Industry**, v. 121, p. 103261, 2020. ISSN 0166-3615. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0166361520304954>>.
- MINKSY, M.; EDMONDS, D. **Stochastic Neural Analogue Reinforcement Calculator**. [S.l.]: Princeton, 1954.
- MIRJALILI, S. Genetic algorithm. Springer, p. 43–55, 2019.
- MITCHELL, M. Genetic algorithms: An overview. v. 1, n. 1, p. 31–39, 1995.
- MOHAČEK-GROŠEV, V.; ĐUROKOVIĆ, M.; MAKSIMOVIĆ, A. Combining raman spectroscopy, dft calculations, and atomic force microscopy in the study of clinker materials. **Materials**, Multidisciplinary Digital Publishing Institute, v. 14, n. 13, p. 3648, 2021.
- NGUYEN, H. T.; WORRING, M.; BOOMGAARD, R. V. D. Watersnakes: Energy-driven watershed segmentation. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 25, n. 3, p. 330–342, 2003.
- NOCK, R.; NIELSEN, F. Statistical region merging. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 26, n. 11, p. 1452–1458, 2004.
- NOZAKI, Y.; WATANABE, T. Development of artificial neural network based automatic stride length estimation method using imu: Validation test with healthy subjects. **IEICE TRANSACTIONS on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 103, n. 9, p. 2027–2031, 2020.
- PARIS, S.; KORNPORST, P.; TUMBLIN, J.; DURAND, F. **Bilateral filtering: Theory and applications**. [S.l.]: Now Publishers Inc, 2009.
- PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V. *et al.* Scikit-learn: Machine learning in python. **Journal of machine learning research**, v. 12, n. Oct, p. 2825–2830, 2011.
- PIEPER, C.; WIRTZ, S.; SCHAEFER, S.; SCHERER, V. Numerical investigation of the impact of coating layers on rdf combustion and clinker properties in rotary cement kilns. **Fuel**, Elsevier, v. 283, p. 118951, 2021.
- PODDER, S.; BHATTACHARJEE, S.; ROY, A. An efficient method of detection of covid-19 using mask r-cnn on chest x-ray images. **AIMS Biophysics**, v. 8, p. 281–290, 07 2021.
- REN, S.; HE, K.; GIRSHICK, R.; SUN, J. Faster r-cnn: Towards real-time object detection with region proposal networks. Curran Associates, Inc., v. 28, 2015. Disponível em: <<https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf>>.

- ROBOFLOW, I. **Roboflow**. 2022. Acessado: 14-11-2022. Disponível em: <<https://app.roboflow.com>>.
- RODRIGUES, D.; JR, A. S.; JR, E. C.; COSTA, A. Modelagem matemática de um forno rotativo empregado na produção de clínquer. **Cerâmica**, SciELO Brasil, v. 59, p. 302–309, 2013.
- ROSASCO, L.; VITO, E. D.; CAPONNETTO, A.; PIANA, M.; VERRI, A. Are loss functions all the same? **Neural computation**, MIT Press, v. 16, n. 5, p. 1063–1076, 2004.
- ROSSUM, G. V.; DRAKE, F. L. **Python 3 Reference Manual**. Scotts Valley, CA: CreateSpace, 2009. ISBN 1441412697.
- SHI, X.; HUANG, G.; HAO, X.; YANG, Y.; LI, Z. A synchronous prediction model based on multi-channel cnn with moving window for coal and electricity consumption in cement calcination process. **Sensors**, v. 21, n. 13, 2021. ISSN 1424-8220.
- SOUZA, V. C. G.; SAMPAIO, C. H.; TAVARES, L. M. M. Estudo por microscopia óptica e lupa das características mineralógicas e microestruturais do clínquer aplicado ao processo de moagem e qualidade do cimento. **Rem: Revista Escola de Minas**, SciELO Brasil, v. 55, p. 125–129, 2002.
- STUTZMAN, P. E.; BULLARD, J. F.; FENG, P.; STUTZMAN, P. E. **Quantitative imaging of clinker and cement microstructure**. [S.l.]: US Department of Commerce, National Institute of Standards and Technology, 2016.
- STUTZMAN, P. E.; CLIFTON, J. R. *et al.* Specimen preparation for scanning electron microscopy. v. 21, p. 10–22, 1999.
- SUTAWIJAYA, A. H.; KAYI, A. *et al.* Optimizing the clinker production by using an automation model in raw material feed. **International Journal of Industrial Optimization**, v. 2, n. 1, p. 17–32, 2021.
- SZUMMER, M.; PICARD, R. W. Indoor-outdoor image classification. In: IEEE. **Proceedings 1998 IEEE International Workshop on Content-Based Access of Image and Video Database**. [S.l.], 1998. p. 42–51.
- TORREY, L.; SHAVLIK, J. Transfer learning. IGI global, p. 242–264, 2010.
- WOLLSCHLAEGER, M.; SAUTER, T.; JASPERNEITE, J. The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0. **IEEE Industrial Electronics Magazine**, v. 11, n. 1, p. 17–27, 2017.
- XIE, F.; GAO, Q.; JIN, C.; ZHAO, F. Hyperspectral image classification based on superpixel pooling convolutional neural network with transfer learning. **Remote Sensing**, v. 13, n. 5, 2021. ISSN 2072-4292.
- XU, J.; FU, D.; SHAO, L.; ZHANG, X.; LIU, G. A soft sensor modeling of cement rotary kiln temperature field based on model-driven and data-driven methods. **IEEE Sensors Journal**, v. 21, n. 24, p. 27632–27639, 2021.
- XU, L. D.; XU, E. L.; LI, L. Industry 4.0: state of the art and future trends. **International Journal of Production Research**, Taylor Francis, v. 56, n. 8, p. 2941–2962, 2018.



ZHENG, X.; LEI, Q.; YAO, R.; GONG, Y.; YIN, Q. Image segmentation based on adaptive k-means algorithm. **EURASIP Journal on Image and Video Processing**, Springer, v. 2018, n. 1, p. 1–10, 2018.